

Please type a plus sign (+) inside this box

+

PTO/SB/05 (2/98)

Approved for use through 09/30/00. OMB 0651-0032  
Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE

Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

UTILITY  
PATENT APPLICATION  
TRANSMITTAL

Only for new nonprovisional applications under 37 CFR 1.53 (b)

Attorney Docket No.	38-21(51376)B
First Named Inventor or Application Identifier	Corbin, David R.
Title	BACILLUS THURINGIENSIS CHROMOSOMAL GENOME SEQUENCES AND USES THEREOF
Express Mail Label No.	EK016827222US

## APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents

## ADDRESS TO:

Assistant Commissioner for Patents  
Box Patent Application  
Washington, DC 202311. ☒ \*Fee Transmittal Form (Form PTO-1082)  
(Submit an original and a duplicate for fee processing)2. ☒ Specification [Total Pages 312]  
(preferred arrangement set forth below)

- Descriptive title of the Invention
- Cross References to Related Applications
- Statement Regarding Fed sponsored R and D
- Reference to Microfiche Appendix
- Background of the Invention
- Brief Summary of the Invention
- Brief Description of the Drawings (if filed)
- Detailed Description
- Claims
- Abstract of the Disclosure

☐ Drawing(s) (35 USC 113) [Total Sheets ]

4. Oath or Declaration [Total Pages ]

- a. ☐ Newly executed (original or copy)
- b. ☐ Copy from a prior application (37 CFR 1.63(d))  
(for continuation/divisional with Box 17 completed)

[Note Box 5 below]

i. ☐ DELETION OF INVENTOR(S)

Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR 1.63(d)(2) and 1.33(b).

5. ☐ Incorporation By Reference (useable if Box 4b is checked)

The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.

6. ☐ Microfiche Computer Program (Appendix)7. Nucleotide and/or Amino Acid Sequence Submission  
(if applicable, all necessary)

- a. ☒ Computer Readable Copy
- b. ☐ Paper Copy (identical to computer copy)
- a. ☐ Statement verifying identity of above copies

## ACCOMPANYING APPLICATION PARTS

8. ☐ Assignment Papers (cover sheet document(s))
9. ☐ 37 CFR 3.73(b) Statement ☐ Power of Attorney  
(when there is an assignee)
10. ☐ English Translation Document (if applicable)
11. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations
12. ☐ Preliminary Amendment
13. ☒ Return Receipt Postcard (MPEP 503) (Two)  
(should be specifically itemized)
14. ☐ \*Small Entity ☐ Statement filed in prior application, status still proper and desired
15. ☐ Certified Copy of Priority Document (s)  
(if foreign priority is claimed)
16. ☐ Other:

\*NOTE FOR ITEMS 1 and 14 IN ORDER TO BE ENTITLED TO PAY SMALL ENTITY FEES, A SMALL ENTITY STATEMENT IS REQUIRED (37 C.F.R. §1.27), EXCEPT IF ONE FILED IN A PRIOR APPLICATION IS RELIED UPON (37 C.F.R. §1.28)

If a CONTINUING APPLICATION, check appropriate box and supply the requisite information:

17. ☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application No:  
Prior Application Information: Examiner: Group/Art Unit:

## 18. CORRESPONDENCE ADDRESS

☐ Customer Number or Bar Code Label or ☒ Correspondence address below  
(Insert Customer No. or attach Bar Code Label here)

NAME	Timothy K. Ball, PhD				
ADDRESS	Patent Department, E2NA; Monsanto Company, 800 N. Lindbergh Boulevard				
CITY	St. Louis	STATE	Missouri	ZIP CODE	63167
COUNTRY	US	TELEPHONE	(314) 737-7387	FAX	(314) 737-6047

Name (Print/Type)	Timothy K. Ball, PhD	Registration No.	42287
Signature	Timothy K. Ball, PhD	Date:	Sept. 15, 2000

Burden Hour Statement This form is estimated to take 2 hours to completed. Time will vary depending on the needs of the individual cases. Any comments on the amount of time you required to completed this form should be sent to the Office of Assistance Quality and Enhancement Division, Patent and Trademark Office, Washington, DC 20231, and to the Office of Information and Regulatory Affairs, Office of Management and Budget (Project 0651-0037), Washington, DC 20503. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO Assistant Commissioner for Patents, Washington, DC 20231

by [Signature] (Jian S. Zhou) Reg. No. 41,422

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

# FEE TRANSMITTAL for FY 1999

Patent fees are subject to annual revision.  
Small Entity payments must be supported by a small entity statement,  
otherwise large entity fees must be paid. See Forms PTO/SB/09-12.  
See 37 C.F.R. §§ 1.27 and 1.28.

TOTAL AMOUNT OF PAYMENT (\$ 1,906.00

## Complete if Known

Application Number  
Filing Date  
First Named Inventor David R. Corbin  
Examiner Name To Be Assigned  
Group / Art Unit  
Attorney Docket No. 38-21(51376)B

09/15/00  
U.S. PTO  
09/663779

## METHOD OF PAYMENT (check one)

1. ☒ The Commissioner is hereby authorized to charge indicated fees and credit any over payments to

Deposit Account Number 13-4125

Deposit Account Name

☒ Charge Any Additional Fee Required  
Under 37 CFR §§ 1.16 and 1.17

2. ☐ Payment Enclosed:  
☐ Check ☐ Money Order ☐ Other

## FEE CALCULATION

### 1. BASIC FILING FEE

Large Entity Fee Code (\$)	Small Entity Fee Code (\$)	Fee Description	Fee Paid
101 760	201 380	Utility filing fee	690.00
106 310	206 155	Design filing fee	
107 480	207 240	Plant filing fee	
108 760	208 380	Reissue filing fee	
114 150	214 75	Provisional filing fee	

SUBTOTAL (1) (\$ 690.00

### 2. EXTRA CLAIM FEES

Total Claims	Extra Claims	Fee from below	Fee Paid
50	-20** = 30	x 18	= 540
Independent Claims	10	-3** = 7	x 78 = 546
Multiple Dependent			= 0

\*\*or number previously paid, if greater; For Reissues, see below

Large Entity Fee Code (\$)	Small Entity Fee Code (\$)	Fee Description
103 18	203 9	Claims in excess of 20
102 78	202 39	Independent claims in excess of 3
104 260	204 130	Multiple dependent claim, if not paid
109 78	209 39	** Reissue independent claims over original patent
110 18	210 9	** Reissue claims in excess of 20 and over original patent

SUBTOTAL (2) (\$ 1,086.00

## FEE CALCULATION (continued)

### 3. ADDITIONAL FEES

Large Entity Fee Code (\$)	Small Entity Fee Code (\$)	Fee Description	Fee Paid
105 130	205 65	Surcharge - late filing fee or oath	0.00
127 50	227 25	Surcharge - late provisional filing fee or cover sheet	0.00
139 130	139 130	Non-English specification	0.00
147 2,520	147 2,520	For filing a request for reexamination	0.00
112 920*	112 920*	Requesting publication of SIR prior to Examiner action	0.00
113 1,840*	113 1,840*	Requesting publication of SIR after Examiner action	0.00
115 110	215 55	Extension for reply within first month	0.00
116 380	216 190	Extension for reply within second month	0.00
117 870	217 435	Extension for reply within third month	0.00
118 1,360	218 680	Extension for reply within fourth month	0.00
128 1,850	228 925	Extension for reply within fifth month	0.00
119 300	219 150	Notice of Appeal	0.00
120 300	220 150	Filing a brief in support of an appeal	0.00
121 260	221 130	Request for oral hearing	0.00
138 1,510	138 1,510	Petition to institute a public use proceeding	0.00
140 110	240 55	Petition to revive - unavoidable	0.00
141 1,210	241 605	Petition to revive - unintentional	0.00
142 1,210	242 605	Utility issue fee (or reissue)	0.00
143 430	243 215	Design issue fee	0.00
144 580	244 290	Plant issue fee	0.00
122 130	122 130	Petitions to the Commissioner	130.00
123 50	123 50	Petitions related to provisional applications	0.00
126 240	126 240	Submission of Information Disclosure Stmt	0.00
581 40	581 40	Recording each patent assignment per property (times number of properties)	0.00
146 760	246 380	Filing a submission after final rejection (37 CFR § 1.129(a))	0.00
149 760	249 380	For each additional invention to be examined (37 CFR § 1.129(b))	0.00
Other fee (specify)			0.00
Other fee (specify)			0.00

\* Reduced by Basic Filing Fee Paid

SUBTOTAL (3) (\$ 130.00

## SUBMITTED BY

Name (Print/Type) Jian S. Zhou Registration No. (Attorney/Agent) 41,422 Telephone 314-694-8908  
Signature Date Sept. 15, 2000

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO Assistant Commissioner for Patents, Washington, DC 20231.

Express Mail Label No. EK016827222US



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

jc925 U.S. PTO  
09/663779  
09/15/00

In re application of:

**David R. Corbin et al.**

Appln. No.: To Be Assigned

Filed:

For: ***Bacillus thuringiensis Chromosomal  
Genome Sequences and Uses thereof***

Art Unit: To Be Assigned

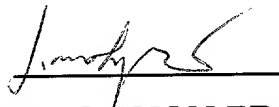
Examiner: To Be Assigned

Atty. Docket: 38-21(51376)B

I hereby verify this transmittal together with the patent application referred to below is being deposited  
with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10  
Express Mail Label No. EK016827222US on the date indicated and is addressed to:

Commissioner for Patents, Washington, D.C. 20231

Jian S. Zhou Registration No. 41, 422



Date: Sept. 15, 2000

**TRANSMITTAL OF NON-PROVISIONAL PATENT APPLICATION**

Commissioner for Patents  
Washington, D.C. 20231

***Box Patent Application***

Sir:

Transmitted herewith for filing under 37 C.F.R. §1.53(b) is the complete patent application  
of :

**BACILLUS THURINGIENSIS CHROMOSOMAL GENOME SEQUENCES AND USES THEREOF**

Inventors: David R. CORBIN  
Thomas M. Malvar  
Hridayabhiranjan Shukla

The following documents are forwarded herewith for appropriate action by the U.S. Patent  
and Trademark Office:

1. Utility Patent Application Transmittal (PTO/SB/05);
2. Form PTO-1082 (in duplicate);

3. U.S. Utility Patent Application entitled:

**BACILLUS THURINGIENSIS CHROMOSOMAL GENOME SEQUENCES  
AND USES THEREOF**


and naming as inventors:

David R. CORBIN, Thomas M. MALVAR, and Hridayabhiranjan SHUKLA  
the application consisting of:

- a. A specification containing:
- (i) 304 pages of a description prior to the claims;
  - (ii) 7 pages of claims (50 claims); and
  - (iii) a one (1) page abstract;
4. A CD-ROM containing the Sequence Listing;
5. Petition to Suspend Sequence Rules (in duplicate); and
6. Two (2) return postcards.

It is respectfully requested that the two attached prepaid postcards be stamped with the filing date and unofficial application number and returned as soon as possible.

Respectfully submitted,



Timothy K. Ball, PhD Esq (Reg. No. 42,287)  
Jian S. Zhou, PhD (Reg. No. 41,422)

Enclosures

Date: Sept. 15, 2000

Patent Department, E2NA  
Monsanto Company  
St. Louis, MO 63167  
Tel: 636-737-7387  
Fax: 636-737-6047

**APPLICATION FOR PATENT**  
**FOR**  
***BACILLUS THURINGIENSIS* CHROMOSOMAL GENOME SEQUENCES AND USES**  
**THEREOF**

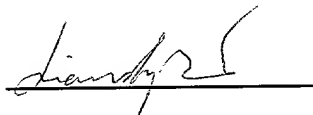
by

**David R. Corbin, Thomas M. Malvar, and Hridayabhiranjan Shukla**

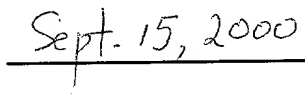
I hereby verify the above-identified patent application is being deposited with  
the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10  
Express Mail Label No. EK016827222US on the date indicated and is addressed to:  
Commissioner for Patents, Washington, D.C. 20231

**Jian S. Zhou**

Registration No. 41, 422



Date:



## CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority under 35 U.S.C §119(e) of U.S. Provisional Application Serial No. 60/154,678 filed on September 17, 1999, the entire content of which is incorporated herein by reference.

5

## FIELD OF THE INVENTION

The present invention relates to nucleic acid sequences from *Bacillus thuringiensis* and, in particular, to chromosomal genomic DNA sequences. The invention encompasses nucleic acid molecules present in non-coding regions as well as nucleic acid molecules that encode proteins and fragments of proteins. Nucleic acid sequences that encode proteins and/or enzymes and homologues and fragments thereof are encompassed by the invention including but not limited to insect inhibitory proteins, proteins capable of conferring antibiotic resistance, microbial inhibitory proteins including bactericidal, bacteriostatic, fungicidal, and fungistatic proteins, polyketide synthases, transposons and mobile genetic elements and their corresponding transposases, excisases and integrases, phage and phage particle proteins, other useful protein homologues, ribosomal RNA (rRNA), and transfer RNA (tRNA). In addition, proteins and fragments thereof so encoded and antibodies capable of binding the proteins are encompassed by the present invention. The invention also relates to methods of using the disclosed nucleic acid molecules, proteins, fragments of proteins, and antibodies, for example, for gene identification and analysis, preparation of constructs, transformation of cells with nucleotide compositions disclosed herein to produce *Bacillus thuringiensis* proteins or fragments thereof, in particular novel insect inhibitory, bactericidal, fungicidal and nematocidal proteins.

## BACKGROUND OF THE INVENTION

*Bacillus thuringiensis* is a spore-forming Gram-positive bacterium. During sporulation, *B. thuringiensis* produces proteinaceous inclusions which are composed of proteins known as insecticidal crystal proteins (ICPs), Cry proteins, or delta-endotoxins. These proteins are toxic to a variety of insect species including orders Lepidoptera, Coleoptera, Diptera, Hemiptera, Hymenoptera, Orthoptera, and Mallophaga and to nematodes, mites, and protozoa (Beegle and Yamamoto, *Can. Entomol.* 124:587-616; Feitelson, *Advanced Engineered Pesticides* (L. Kim, *ed.*), Marcel Dekker, Inc., New York (1993), pp. 63-71; Feitelson, *et al.*, *Bio/Technology* 10:271-275; U.S. Patent No. 4,948,734 (1990)). Due to their high specificity for particular insect pests and their safety for man and the environment, ICPs have been used as biopesticides for the last three decades. Using molecular genetic techniques, numerous delta-endotoxin

genes have been isolated and their DNA sequences determined. The cloning and sequencing of a number of  $\delta$ -endotoxin genes from a variety of *B. thuringiensis* strains has been described and are summarized by Schnepf *et al.* (*Microbiol. Mol. Biol. Rev.* 62:775-806, *Bacillus thuringiensis* And Its Pesticidal Crystal Proteins, 1998). The nomenclature and appearance of newly identified genes is summarized and

5 regularly updated at [http://www.biols.susx.ac.uk/Home/Neil\\_Crickmore/Bt/](http://www.biols.susx.ac.uk/Home/Neil_Crickmore/Bt/). These genes have been used to develop certain genetically engineered *B. thuringiensis* products that are in commercial use. Recent developments have seen new  $\delta$ -endotoxin delivery systems developed, including genetically engineered plants that contain and express  $\delta$ -endotoxin genes. *Bacillus thuringiensis* is a key source of genes, which when modified can be used for transgenic expression to provide pest resistance in plants.

10 *B. thuringiensis* strains are classified into subspecies or varieties, based on biochemical and serological criteria (de Barjac, *Entomophaga* 7: 5-61 (1962); de Barjac, *Proceedings of the IIIrd International Colloquium on Invertebrate Pathology* (C.C. Payne and H.D. Burges, eds.), Society for Insect Pathology, U.K., 451-453 (1982)). Each subspecies may produce one or several insecticidal protein toxins. To date, approximately 172  $\delta$ -endotoxins belonging to 28 classes have been identified.

15 There is also a nonprotein toxin, the  $\beta$ -exotoxin, secreted by some *B. thuringiensis* strains. This toxin, which is assayed on house fly larvae (Sêbesta *et al.*, "Thuringiensin, the  $\beta$ -exotoxin of *Bacillus thuringiensis*," in W.H. Burgess (ed.), *Microbial Control of Pests and Plant Diseases, 1970-1980*, Academic Press, Inc., New York, pp. 249-281 (1981)), is not as selective as the  $\delta$ -endotoxins.

20 Extensive studies have been carried out with *B. thuringiensis* subspecies that produce proteinaceous inclusions during sporulation. The inclusions are often bipyramidal, but some are cuboidal or multifaceted, and there is a wide variety of other morphologies. Some strains contain more than one type of inclusion in each cell. These inclusions are present within the mother cell adjacent to the spore, but in a few subspecies, they are localized within the exosporium (Aronson *et al.*, *Bacteriol. Rev.* 40:360-402 (1976)). Inclusions are released, as is the spore, upon cell lysis.

25 *Bacillus* strains can have a chromosomal genome size of 2.4 to 5.7 Mbp (Carlson, *et al.*, *Appl. Environ. Microbiol.* 60: 1719-1725 (1994)). Physical maps of chromosomes of two *B. thuringiensis* strains, *B. thuringiensis* subsp. *Berliner* 1715 and *B. thuringiensis* subsp. *Thuringiensis* HD2, have been constructed and are estimated to be between 5.4 and 5.7 Mbp (Carlson, *et al.*, *Microbiol.* 142: 1625-1634 (1996); Carlson and Kolstø, *J. Bacteriol.* 175: 1053-1060 (1993)). The total genomes of each of these two

30 strains consist of one or more chromosomes, and a more variable component comprised of extrachromosomal elements (Carlson and Kolstø, *Mol. Microbiol.* 13:161-169 (1994)).

Most *B. thuringiensis* isolates have several extrachromosomal elements, some of them circular plasmids and others linear (Carlson, *et al.*, *Microbiol.* 60: 1719-1725 (1994)). In general, crystal-protein genes are localized on large plasmids (ca. 40 to 200 Mda) of *B. thuringiensis* (Gonzalez, *et al.*, *Plasmid* 5: 351-365 (1981); Carlton and Gonzalez, *Molecular Biology of Microbial Differentiation*, American Society for Microbiology, Washington, D.C. 246-252 (1985), Kronstad, *et al.*, *J Bacteriol.* 154: 419-428 (1983)), and in some cases, more than one gene is present on a given plasmid (Aronson *et al.*, *Bacteriol. Rev.* 40:360-402 (1976); Carlton *et al.*, "The genetics and molecular biology of *Bacillus thuringiensis*," in D.A. Dubnau (ed.), *The Molecular Biology of the Bacilli*, Vol. II, Academic Press, Inc., New York, pp. 211-249 (1985)). However, chromosomal crystal-protein genes have been reported in some *B. thuringiensis* strains (Carlson and Kolstø, *J. Bacteriol.* 175: 1053-1060 (1993), Klier, *et al.*, *EMBO J* 1: 791-799 (1982), Kronstad, *et al.*, *J Bacteriol.* 154: 419-428 (1983)).

*Bacillus thuringiensis* strains often contain multiple epigenetic elements which are known to harbor genes expressing vegetative insecticidal proteins (VIP's) and Bt crystalline insecticidal and nematocidal proteins. It is believed that many other Bt insecticidal / nematocidal genes are present within these sequences, some of which may only be expressed under conditions which cannot be artificially simulated, some of which may be cryptic, and some of which may be actively expressed but which have not been previously identified due to their limited availability as a result of very low levels of expression. Identification of whole or substantial portions of DNA sequences of individual plasmids would greatly facilitate identification of genes encoding novel insect inhibitory proteins. However, when one tries to isolate and purify plasmid DNA of a *B. thuringiensis* species for constructing genomic DNA libraries used in sequencing, it would be difficult to eliminate the contamination of chromosomal DNA. Such contamination would complicate greatly the sequencing effort of individual plasmids and subsequently hinder construction of genetic maps of individual plasmids of the *B. thuringiensis* species. Thus, it would be desirable to generate the complete DNA sequence of the chromosomal genome exclusive of epigenetic sequences of a *B. thuringiensis* species, because the complete DNA sequence of the chromosome could be used as a background to significantly minimize the interference of chromosomal DNA sequences in identification of whole or a substantial portion of individual plasmids and of novel genes encoding insect inhibitory proteins.

Furthermore, although it is unexpected that the complete DNA sequence of the *Bacillus thuringiensis* chromosomal genome exclusive of epigenetic sequences would provide a substantial number of Bt crystalline insecticidal / nematocidal and VIP genes for second generation insect / pest control in crop species, comparison of the open reading frames present within the *Bacillus thuringiensis* chromosomal genome with other bacterial genome sequences, in particular other *Bacillus* species

genomic sequences would allow the subtraction of common sequences and thus the identification of sequences novel and unique to *Bacillus thuringiensis*, and which may play a role in the regulation of expression or activity of genes encoding insecticidal proteins, and may also provide a plethora of useful genes for future insect resistance management technologies and applications. Therefore, it is  
 5 advantageous to generate the complete DNA sequence of the chromosomal genome exclusive of epigenetic sequences of a *B. thuringiensis* species.

Chromosomal genome sequence information from *B. thuringiensis* allows comparisons of those sequences with sequences from other *B. thuringiensis* strains as well as comparisons with DNA sequences from other organisms, including plants, mammals such as humans, bacteria, and fungi such as  
 10 yeasts. In addition, genome sequencing and mapping provides increased opportunities for identification and isolation of agents of commercial interest, as well as insight into mechanisms of genome interactions.

### SUMMARY OF THE INVENTION

The present invention provides an isolated and purified nucleic acid molecule having a first  
 15 nucleotide sequence, wherein: (1) the first nucleotide sequence hybridizes under stringent conditions to a second nucleotide sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283 or complement thereof, wherein the hybridizing portion of the second nucleotide sequence is at least 50 nucleotides in length; (2) the first nucleotide sequence is a portion of third nucleotide sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283; or (3) the first nucleotide  
 20 sequence is the complement of (1) or (2).

In a preferred embodiment, the hybridizing portion of the second nucleotide sequence is at least 100 nucleotides in length. In a more preferred embodiment, the hybridizing portion of the second nucleotide sequence is at least 200 nucleotides in length. In a further more preferred embodiment, the hybridizing portion of the second nucleotide sequence encodes any polypeptide or protein or set forth in  
 25 Table 1.

The present invention also provides an isolated and purified nucleic acid molecule comprising a nucleotide sequence, wherein: (1) the nucleotide sequence encodes any polypeptide or protein set forth in Table 1; or (2) the nucleotide sequence is the complement of (1).

The present invention, in another aspect, provides a substantially purified polypeptide or protein  
 30 comprising an amino acid sequence, wherein the amino acid sequence is defined as follows: (1) the amino acid sequence is encoded by a first nucleotide sequence which specifically hybridizes to the complement of a second nucleotide sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283; or (2) the amino acid sequence is encoded by a third nucleotide sequence

that is at least 50% identical to a portion of the complement of a sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283. In alternative embodiments, the above described third nucleotide sequence is at least 55% identical, at least 60% identical, at least 65% identical, at least 70% identical, at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, or at least 95% identical to a portion of a sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283; and, the above described third nucleotide sequence is identical to a portion of a sequence selected from SEQ ID NO: 1 through SEQ ID NO: 8283.

The present invention also provides a recombinant construct comprising: (A) a promoter region which functions in a host cell to cause the production of an mRNA molecule; which is operably linked to (B) a structural nucleotide sequence, wherein the structural nucleotide sequence encodes a polypeptide or protein set forth in Table 1; which is operably linked to (C) a 3' non-translated sequence that functions in said cell to cause termination of transcription.

The present invention also provides a recombinant construct comprising: (A) a promoter region which functions in a host cell to cause the production of an mRNA molecule wherein the promoter region is selected from the group consisting of promoter sequences located within SEQ ID NO: 1 through SEQ ID NO: 8283 or complements thereof; which is linked to (B) a structural nucleotide sequence encoding a polypeptide; which is linked to (C) a 3' non-translated sequence that functions in said cell to cause termination of transcription.

The present invention also provides a transformed cell having an exogenous nucleic acid molecule which comprises: (A) a promoter region which functions in said cell to cause the production of an mRNA molecule; which is operably linked to (B) a structural nucleic acid molecule, wherein the structural nucleotide encodes any polypeptide or protein set forth in Table 1; which is operably linked to (C) a 3' non-translated sequence that functions in said cell to cause termination of transcription.

The present invention also provides a transformed cell having an exogenous nucleic acid molecule which comprises: (A) a promoter region which functions in said cell to cause the production of an mRNA molecule wherein the promoter region is selected from the group consisting of promoter sequences located within SEQ ID NO: 1 through SEQ ID NO: 8283 or complements thereof; which is operably linked to (B) a structural nucleotide sequence encoding a polypeptide; which is operably linked to (C) a 3' non-translated sequence that functions in said cell to cause termination of transcription.

The present invention also provides a plant cell, a mammalian cell, a bacterial cell, an algal cell, an insect cell and a fungal cell transformed with an isolated nucleic acid molecule of the present invention.



The invention also provides isolated nucleic acid molecules comprising nucleotide sequences encoding polypeptides or proteins exhibiting insect inhibitory activity, wherein said activity is manifested by inhibiting the growth or development of, or contributing substantially to, or causing the death of a Coleopteran, a Dipteran, a Lepidopteran, a Hemipteran, a Hymenopteran, or a sucking and piercing insect or insect larvae thereof. Also provided are nucleotide sequences encoding novel proteins comprising polypeptides which augment the activity of polypeptides exhibiting insect inhibitory activity when fed to Coleopteran, Dipteran, Lepidopteran, Hemipteran, Hymenopteran, or sucking and piercing insects or insect larvae thereof.

The present invention also provides a method for using insect inhibitory proteins for controlling target insect pests, i.e. also known as insect pest control.

The present invention also provides a computer readable medium having recorded thereon one or more nucleotide sequences, wherein each of the nucleotide sequences is selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283 or complements thereof.

The present invention also provides a computer readable medium having recorded thereon one or more of the nucleotide sequences encoding a protein or fragment thereof set forth in Table 1.

The present invention also provides a method for using the computer media of the present invention in isolating/identifying nucleic acids encoding insect inhibitory proteins, or proteins involved in biosynthesis of antibiotics.

The present invention also provides a method for identifying one or more genes encoding insect inhibitory proteins in the sequences of one or more plasmids of a *Bacillus thuringiensis*, wherein the method of the present invention comprises the steps of: a) isolating and purifying plasmid DNA; b) constructing a DNA library from the isolated and purified plasmid DNA; c) sequencing the DNA library to obtain a set of plasmid DNA sequences; d) comparing the set of DNA sequences with a set of chromosomal DNA sequences, wherein the set of chromosomal DNA sequences comprises the group consisting of SEQ ID No: 1 through SEQ ID No: 8283; e) identify common sequences, which are identified both in the set of plasmid DNA sequences and in the set of chromosomal DNA sequences; f) subtracting the common sequences from the set of plasmid DNA sequences to obtain a subtracted set of plasmid DNA sequences; g) assembling the subtracted set of DNA sequences to obtain contigs and sequence assemblies; h) determining open reading frames in the contigs and sequence assemblies; and h) identifying one or more genes encoding insect inhibitory proteins in the sequences of one or more plasmids of said *Bacillus thuringiensis*.

The present invention also provides a method for identifying plasmid DNA sequences of a *Bacillus* species, the method comprising the steps of a) identifying a *Bacillus* species strain which does

not contain plasmid DNA; b) generating a library of chromosomal genomic DNA from said *Bacillus* species strain which does not contain plasmid DNA; c) obtaining the nucleotide sequence of said chromosomal genomic DNA; d) identifying a *Bacillus* species strain which contains plasmid DNA; e) generating a library of said *Bacillus* species plasmid DNA; f) obtaining the nucleotide sequence of said plasmid DNA; g) subtracting any common sequences identified in the plasmid DNA which are also identified in the chromosomal genomic DNA; and h) constructing a contig sequence or sequence assemblies of said plasmid DNA, wherein said contig sequence or sequence assemblies comprise the plasmid DNA sequence of said *Bacillus* species.

## DETAILED DESCRIPTION OF THE INVENTION

### Agents of the Present Invention

#### Nucleic Acid Molecules

One aspect of the present invention relates to an isolated nucleic acid molecule having a first nucleotide sequence, wherein: (1) the first nucleotide sequence hybridizes under stringent conditions to a second nucleotide sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283 or complements thereof, wherein the hybridizing portion of the second nucleotide sequence is at least 50 nucleotides in length; (2) the first nucleotide sequence is a portion of any sequence selected from the group consisting of SEQ ID NO:1 through SEQ ID NO: 8283; or (3) the first nucleotide sequence is the complement of (1) or (2).

The term "nucleic acid" means a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. Nucleic acids may also optionally contain synthetic, non-natural or altered nucleotide bases that permit correct read through by a polymerase and do not alter expression of a polypeptide encoded by that nucleic acid.

The term "an isolated nucleic acid" refers to a nucleic acid that is no longer accompanied by some of materials with which it is associated in its natural state or to a nucleic acid the structure of which is not identical to that of any of naturally occurring nucleic acid. Examples of an isolated nucleic acid include but are not limited to: (1) DNAs which have the sequence of part of a naturally occurring genomic DNA molecules but are not flanked by two coding sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs; (2) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (3) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment;

(4) recombinant DNAs; and (5) synthetic DNAs. An isolated nucleic acid may also be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

It is also contemplated by the inventors that the isolated nucleic acids of the present invention may also include known types of modifications, for example, labels which are known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog. Other known modifications include inter-nucleotide modifications, for example, those with uncharged linkages (methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.) and with charged linkages (phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, proteins (including nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), those with intercalators (acridine, psoralen, etc.), those containing chelators (metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, and those with modified linkages.

The term "nucleotide sequence" refers to both the sense and antisense strands of a nucleic acid as either individual single strands or in the duplex. It includes, but is not limited to, self-replicating plasmids, chromosomal sequences, and infectious polymers of DNA or RNA.

A nucleotide sequence is said to be the "complement" of another nucleotide sequence if they exhibit complete complementarity. As used herein, molecules are said to exhibit "complete complementarity" when every nucleotide of one of the sequences is complementary to a nucleotide of the other.

A "coding sequence" or "structural nucleotide sequence" is a nucleotide sequence which is translated into a polypeptide, usually via mRNA, when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the 5'-terminus and a translation stop codon at the 3'-terminus. A coding sequence can include, but is not limited to, genomic DNA, cDNA, and recombinant polynucleotide sequences.

The term "recombinant DNAs" refers to DNAs that contains a genetically engineered modification through manipulation via mutagenesis, restriction enzymes, and the like.

The term "synthetic DNAs" refers to DNAs assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form DNA segments which are then enzymatically assembled to construct the entire DNA. "Chemically synthesized", as related to a sequence of DNA, means that the component nucleotides were assembled in vitro. Manual chemical synthesis of DNA may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines.

The term "stringent conditions" or "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target sequence, to a detectable greater degree than other sequences (e.g., at least 2 fold over background). Stringent conditions are sequence dependent and will be different in different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences can be identified which are 100% complementary to the probe (homologous probing). Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of complementary to the probe are detected (heterologous probing).

Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M monovalent cation (e.g.,  $\text{Na}^+$ ), typically about 0.01 to 1.0 M monovalent cation concentration at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulfate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1 X SSC at 55 to 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1 X SSC at 60 to 65°C.

Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the  $T_m$  can be approximated from the equation of Meinkoth and Wahl, *Anal. Biochem.*, 138:267-284 (1984):  $T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41 (\%GC) - 0.61 (\% \text{ form}) - 500/L$ ; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe.  $T_m$  is reduced by about 1°C for each 1 % of mismatching; thus,  $T_m$  hybridization and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with >90% identity are sought, the  $T_m$  can be decreased 10°C. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point ( $T_m$ ) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4°C lower than the thermal melting point ( $T_m$ ); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point ( $T_m$ ); low stringency conditions can utilize a hybridization

and/or wash at 11, 12, 13, 14, 15, or 20°C lower than the thermal melting point ( $T_m$ ). Using the equation, hybridization and wash compositions, and desired  $T_m$  those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a  $T_m$  of less than 45°C (aqueous solution) or 32°C (formamide solution)

5 it is preferred to increase the SSC concentration so that a higher temperature can be used. Appropriate stringent conditions are known to those skilled in the art or can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6.

The hybridization portion of the two hybridizing nucleic acids is usually at least 40 nucleotides in length, more usually at least about 75 nucleotides in length, more particularly at least 100 nucleotides in  
10 lengths. The hybridizing portion of the hybridizing nucleic acid is at least 80%, at least 90%, or at least 98% identical to the sequence of a portion of a sequence set forth in SEQ ID NO: 1 to SEQ ID NO: 8283.

Another aspect of the present invention relates to an isolated nucleic acid molecule comprising one or more open reading frames listed in Table 1. An "open reading frame" (ORF) is a region of a nucleotide sequence which encodes a polypeptide. This region may represent a portion of a coding  
15 sequence or a total coding sequence. Table 1 sets forth a list of open reading frames identified in the isolated nucleic acid molecules, wherein the open reading frames encode *Bacillus thuringiensis* proteins or polypeptide or fragments thereof which are homologues of known proteins or unknown proteins, or of tRNA's or rRNA's or fragments thereof which are homologues of known tRNA's or rRNA's.

Open reading frames in genomic sequences can be screened for the presence of protein  
20 homologues utilizing a number of different search algorithms that have been developed, one example of which is the suite of programs referred to as BLAST programs. There are five implementations of BLAST, three designed for nucleotide sequences queries (BLASTN, BLASTX, and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, *Trends in Biotechnology* 12:76-80 (1994); Birren *et al.*, *Genome Analysis* 1:543-559 (1997)). Other examples of suitable programs  
25 that can be utilized are well known in the art. In addition, unidentified reading frames may be screened for by gene prediction software such as GenScan, which is located at <http://gnomic.stanford.edu/GENSCANW.html>. Novel genes, i.e., with no known homologs, can be predicted with the program GeneMark, which calculates the probability of a gene based on the presence of a gene-like 'grammar' in the DNA sequence (i.e., start and stop signals, and a significant open reading  
30 frame) and statistical analyses of protein-coding potential through biases in putative codon usage (see <http://genemark.biology.gatech.edu/GeneMark> for details).

The present invention also provides an isolated nucleic acid molecule comprising a first nucleotide sequence, wherein: (1) the first nucleotide sequence hybridizes under stringent conditions to a

second nucleotide sequence, wherein the hybridizing portion of the second nucleotide sequence encodes any polypeptide or protein set forth in Table 1; or (2) the first nucleotide sequence is the complement of (1) or (2).

In one embodiment, an isolated nucleic acid molecule comprises a nucleotide sequence, wherein the nucleotide sequence encodes any polypeptide or protein set forth in Table 1.

The term "polypeptide" or "protein" refers to a linear polymer composed of amino acids connected by peptide bonds.

By "substantial identical" or "substantially identical" as used in reference to two amino acid sequences, it is meant that one amino acid sequence is identical to the other amino acid sequence or has at least 50% sequence identity, at least 70% sequence identity, preferably at least 80%, more preferably at least 90%, and most preferably at least 95% identity when compared to the other amino acid sequence as a reference sequence using the programs described herein; preferably BLAST using standard parameters, as described below.

"Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

Polypeptides which are "substantially similar" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. "Conservative amino acid substitutions" refer to substitutions of one or more amino acids in a native amino acid sequence with another amino acid(s) having similar side chains, resulting in a silent change. Conserved substitutes for an amino acid within a native amino acid sequence can be selected from other members of the group to which the naturally occurring amino acid belongs. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and

methionine. Preferred conservative amino acids substitution groups are: valine-leucine, valine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, aspartic acid-glutamic acid, and asparagine-glutamine.

Optimal alignment of sequences for comparison can use any means to analyze sequence identity (homology) known in the art, e.g., by the progressive alignment method of termed "PILEUP" (Morrison, Mol. Biol. Evol. 14:428-441 (1997), as an example of the use of PILEUP); by the local homology algorithm of Smith & Waterman (Adv. Appl. Math. 2: 482 (1981)); by the homology alignment algorithm of Needleman & Wunsch (J. Mol. Biol. 48:443 (1970)); by the search for similarity method of Pearson (Proc. Natl. Acad. Sci. USA 85: 2444 (1988)); by computerized implementations of these algorithms (e.g., GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI); ClustalW (CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, California, described by, e.g., Higgins, Gene 73: 237-244 (1988); Corpet, Nucleic Acids Res. 16:10881-10890 (1988); Huang, Computer Applications in the Biosciences 8:155-165 (1992); and Pearson, Methods in Mol. Biol. 24:307-331 (1994); Pfam (Sonnhammer, Nucleic Acids Res. 26:322-325 (1998); TreeAlign (Hein, Methods Mol. Biol. 25:349-364 (1994); MES-ALIGN, and SAM sequence alignment computer programs; or, by manual visual inspection.

Another example of algorithm that is suitable for determining sequence similarity is the BLAST algorithm, which is described in Altschul et al, J. Mol. Biol. 215: 403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (NCBI), <http://www.ncbi.nlm.nih.gov/>; see also Zhang, Genome Res. 7:649-656 (1997) for the "PowerBLAST" variation. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al, J. Mol. Biol. 215: 403-410 (1990)). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a wordlength (W) of 11, the BLOSUM62 scoring matrix (see Henikoff, Proc. Natl. Acad. Sci. USA 89:10915-10919(1992)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The term BLAST refers to the BLAST algorithm which performs a statistical analysis of the similarity between two sequences; see, e.g., Karlin, Proc. Natl. Acad. Sci. USA 90:5873-5787 (1993). One measure of similarity provided by the BLAST algorithm is the smallest sum probability ( $P(N)$ ), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

One skilled in the art will recognize that these values of sequence identity can be appropriately adjusted to determine corresponding sequence identity of two nucleotide sequences encoding the proteins of the present invention by taking into account codon degeneracy, conservative amino acid substitutions, reading frame positioning and the like. Substantial identity of nucleotide sequences for these purposes normally means sequence identity of at least 40%, preferably at least 60%, more preferably at least 90%, and most preferably at least 95%.

The term "codon degeneracy" refers to divergence in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a gene for expression in a host cell, it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

The present invention also includes an isolated nucleic acid comprising a nucleotide sequence encoding a polypeptide set forth in Table 1, wherein the an amino acid sequence of the protein or polypeptide contains one or more conservative amino acid substitutions.

In a preferred embodiment of the present invention, the isolated nucleic acid molecule comprising a nucleotide sequence which encodes an insect inhibitory protein or polypeptide or fragment thereof.

In a preferred embodiment of the present invention, the isolated nucleic acid molecule comprising a nucleotide sequence encoding all or substantial portion of a sigma factor homologue listed in Table 2.

The term "insect inhibitory protein" refers to any polypeptide or protein or a substantial portion thereof that exhibits insect inhibitory activity, wherein said activity is manifested by inhibiting the growth or development of, or contributing substantially to, or causing the death of a Coleopteran, a Dipteran, a Lepidopteran, a Hemipteran, a Hymenopteran, or a sucking and piercing insect or insect larvae thereof.



The term "insect inhibitory protein" also refers to any polypeptide or protein with modified amino acid sequence, such as sequence which has been mutated, truncated, increased and the like and which maintains at least the insect inhibitory activity associated with the native protein. Accordingly, the isolated nucleic acids encoding those polypeptide or protein with such modification are also within the scope of the present invention.

In a preferred embodiment of the present invention, the isolated nucleic acid molecule comprising a nucleotide sequence which encodes whole or a portion of a protein homologue capable of conferring antibiotic resistance or resistance to heavy metals or other chemicals, wherein the nucleotide sequence is any sequence set forth in SEQ ID NOs: 33, 98, 145, 162, 180, 204, 275, 298, 361, 397, 421, 423, 579, 613, 624, 692, 726, 862, 930, 950, 986, 995, 1005, 1023, 1130, 1188, 1190, 1208, 1226, 1227, 1240, 1246, 1246, 1257, 1272, 1302, 1339, 1355, 1374, 1393, 1426, 1460, 1471, 1526, 1854, 1914, 1923, 2151, 2179, 2211, and 2304 (Table 4).

In a preferred embodiment of the present invention, the isolated nucleic acid molecule comprising a nucleotide sequence which encodes whole or a portion of a transposon or transposase homologue, wherein the nucleotide sequence is any sequence set forth in SEQ ID Nos: 2, 64, 226, 379, 383, 387, 410, 416, 546, 555, 603, 642, 644, 660, 691, 691, 781, 799, 980, 1002, 1045, 1072, 1098, 1190, 1207, 1214, 1252, 1273, 1275, 1305, 1317, 1330, 1340, 1353, 1354, 1362, 1378, 1378, 1380, 1383, 1386, 1386, 1388, 1391, 1392, 1549, 1573, 1611, 1698, 1725, 1739, 1804, 1869, 1902, 1965, 2041, 2049, 2130, 2135, 2153, 102, 1340, 1795, 1797, 1989, 2055, 2057, 2248, 14, 296, 722, 834, 834, and 999 (Table 3).

In a preferred embodiment of the present invention, the isolated nucleic acid molecule comprising a nucleotide sequence encodes whole or a portion of a toxin or toxin homologue listed in Table 5.

In a preferred embodiment of the present invention, the isolated nucleic acid molecule encodes a *B. thuringiensis* protein or fragment thereof that is a homologue of another *Bacillus* protein. In another preferred embodiment of the present invention, the nucleic acid molecule encodes a *B. thuringiensis* protein or fragment thereof that is a homologue of a fungal protein. In another preferred embodiment of the present invention, the nucleic acid molecule encodes a *B. thuringiensis* protein or fragment thereof that is a homologue of a plant protein. In another preferred embodiment of the present invention, the nucleic acid molecule encodes a *B. thuringiensis* protein or fragment thereof that is a homologue of mammalian protein. In another preferred embodiment of the present invention, a *B. thuringiensis* protein or fragment thereof of the present invention is a homologue of a non-*Bacillus* bacterial protein. In another preferred embodiment of the present invention, the nucleic acid molecule encodes a *B. thuringiensis* protein or fragment thereof that is a homologue of an algal protein.

In a preferred embodiment of the present invention, the nucleic acid molecule of the present invention encodes a *B. thuringiensis* homologue protein or fragment thereof where the *B. thuringiensis* homologue protein exhibits a BLASTP probability score of greater than 1E-12, preferably a BLASTP probability score of between about 1E-30 and about 1E-12, even more preferably a BLASTP probability score of greater than 1E-30 with its homologue. In a preferred embodiment of the present invention, the nucleic acid molecule of the present invention encodes a *B. thuringiensis* homologue protein or fragment thereof exhibits an aa\_tap score of less than 200, preferably an aa\_tap score of between about 200 to about 400, even more preferably an aa\_tap score of greater than 400.

It is understood that certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Because it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence and, of course, its underlying DNA coding sequence and, nevertheless, obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the proteins or fragments of the present invention, or corresponding DNA sequences that encode said peptides, without appreciable loss of their biological utility or activity. It is understood that codons capable of coding for such amino acid changes are known in the art.

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte and Doolittle, *J. Mol. Biol.* 157, 105-132 (1982)). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like.

Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982); these are isoleucine (+4.5), valine (+4.2), leucine (+3.8), phenylalanine (+2.8), cysteine/cystine (+2.5), methionine (+1.9), alanine (+1.8), glycine (-0.4), threonine (-0.7), serine (-0.8), tryptophan (-0.9), tyrosine (-1.3), proline (-1.6), histidine (-3.2), glutamate (-3.5), glutamine (-3.5), aspartate (-3.5), asparagine (-3.5), lysine (-3.9), and arginine (-4.5).

In making such changes, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  is preferred, those which are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Patent 4,554,101, incorporated herein by reference in its entirety, states that the greatest local average hydrophilicity of a protein, as govern by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

5 As detailed in U.S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0), lysine (+3.0), aspartate (+3.0 $\pm$ 1), glutamate (+3.0 $\pm$ 1), serine (+0.3), asparagine (+0.2), glutamine (+0.2), glycine (0), threonine (-0.4), proline (-0.5 $\pm$ 1), alanine (-0.5), histidine (-0.5), cysteine (-1.0), methionine (-1.3), valine (-1.5), leucine (-1.8), isoleucine (-1.8), tyrosine (-2.3), phenylalanine (-2.5), and tryptophan (-3.4).

10 In making such changes, the substitution of amino acids whose hydrophilicity values are within  $\pm 2$  is preferred, those which are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

Another aspect of the present invention relates to a class of isolated nucleic acid molecules comprising promoter sequences or regulatory elements, particularly those found within SEQ ID NO: 1  
15 through SEQ ID NO: 8283 or complements thereof.

The term "promoter sequence" means a nucleotide sequence that is capable of, when located in *cis* to a structural nucleotide sequence encoding a polypeptide or protein, functioning in a way that directs expression of one or more mRNA molecules that encodes the polypeptide or protein. Such promoter regions are typically found upstream of the trinucleotide ATG sequence at the start site of a  
20 protein coding region. Promoter sequences can also include sequences from which transcription of transfer RNA (tRNA) or ribosomal RNA (rRNA) sequences are initiated. Transcription involves the synthesis of an RNA chain representing one strand of a DNA duplex. By "representing" it is meant that the RNA is identical in sequence with one strand of the DNA; it is complementary to the other DNA strand, which provides the template for its synthesis. Transcription takes place by the usual process of  
25 complementary base pairing, catalyzed and scrutinized by the enzyme RNA polymerase. The reaction can be divided into three stages described as initiation, elongation and termination. Initiation begins with the binding of RNA polymerase to the double stranded (DS or ds) DNA. The sequence of DNA required for the initiation reaction defines the promoter. The site at which the first nucleotide is incorporated is called the startsite or startpoint of transcription. Elongation describes the phase during which the enzyme  
30 moves along the DNA and extends the growing RNA chain. Elongation involves the disruption of the DNA double stranded structure in which a transiently unwound region exists as a hybrid RNA-DNA duplex and a displaced single strand of DNA. Termination involves recognition of the point at which no further bases should be added to the chain. To terminate transcription, the formation of phosphodiester

bonds must cease and the transcription complex must come apart. When the last base is added to the RNA chain, the RNA-DNA hybrid is disrupted, the DNA reforms into a duplex state, and the RNA polymerase enzyme and RNA molecule are both released from the DNA. The sequence of DNA required for the termination reaction is called the terminator.

5 Generally, for bacteria the optimal promoter is a sequence consisting of a -35 hexamer separated by about 17 base pairs from a -10 hexamer and lies from about 7 to about 10 base pairs upstream of the startpoint of transcription, but these sequences can vary among and between sequences which are recognized by the RNA polymerase. The startpoint of transcription generally lies from about 20 to about 50 base pairs upstream of the startpoint of translation of one or more open reading frames which  
10 comprise the entire length of an mRNA transcript. Some promoters can be recognized by RNA polymerase alone and in these cases, an accessible promoter will always be transcribed. Promoter availability may be determined by extraneous proteins, which either may act directly at the promoter to block access by RNA polymerase, or may function indirectly by controlling the structure of the genome in the region. Other promoters are not by themselves adequate to support transcription initiation and thus  
15 ancillary protein and or RNA factors are required to further initiation. The additional protein or RNA factors usually act by recognizing sequences of DNA that are close to, or overlap with, the sequence bound by RNA polymerase itself. Additionally, some of these ancillary factors must touch and concern the RNA polymerase in order to effect efficient transcription initiation as well as transcription elongation.

Promoters in particular in *Bacillus* species are highly regulated by the appearance and  
20 disappearance of accessory factors known as sigma factors which touch and concern both the DNA sequences flanking the promoter site as well as the RNA polymerase in order to effect efficient transcription initiation and elongation. Such factors are required for effecting transcription from various classes of promoters and along with other factors expressed from *spoO* genes affect both the temporal regulation of expression from promoters as well as the spatial distribution of patterns of expression within  
25 the cell during differentiation and development of the *Bacillus* cell from spore activation, vegetative growth and proliferation, and sporulation. Examples of sigma factors which are known to function in this manner are  $\sigma^A$ ,  $\sigma^B$ ,  $\sigma^C$ ,  $\sigma^D$ ,  $\sigma^E$ ,  $\sigma^F$ ,  $\sigma^G$ ,  $\sigma^H$ ,  $\sigma^I$ , and  $\sigma^K$ . In addition, the sporulation factors which function for temporal and spatial regulatory gene expression include gene products in the classes of *SpoO*, *SpoOI*, and *SpoOII*.

30 Thus, in *Bacillus*, in many circumstances a consensus promoter may be sufficient for expression, however additional sequences may act to further regulate expression, for example, in response to biochemical, developmental or environmental signals. In a preferred embodiment of the present

invention, the promoter is upstream of a nucleic acid sequence that encodes a *B. thuringiensis* protein or fragment thereof.

Promoters of the present invention can be included within sequences up to 10 kb upstream of the trinucleotide ATG sequence at the start site of a protein coding region, tRNA, or rRNA. Promoters of the present invention can preferably be included within sequences up to 5 kb upstream of the trinucleotide ATG sequence at the start site of a protein coding region, tRNA or rRNA. Promoters of the present invention can more preferably be included within sequences up to 2kb upstream of the trinucleotide ATG sequence at the start site of a protein coding region, tRNA or rRNA. Promoters of the present invention can most preferably be included within sequences up to 500 bp upstream of the trinucleotide ATG sequence at the start site of a protein, tRNA, or rRNA coding region. While in many circumstances a 300 bp promoter may be sufficient for expression, additional sequences may act to further regulate expression, for example, in response to biochemical, developmental or environmental signals. In a preferred embodiment of the present invention, the promoter is upstream of an nucleic acid sequence that encodes a *Bacillus thuringiensis* protein or fragment thereof.

The term "regulatory element" is intended to mean a series of nucleotides that determines if, when, and at what level a particular gene is expressed. Regulatory DNA sequences specifically interact with regulatory or other proteins. Many regulatory elements act in *cis* ("*cis* elements") and are believed to affect DNA topology, producing local conformations that selectively allow or restrict access of RNA polymerase to the DNA template or that facilitate selective opening of the double helix at the site of transcriptional initiation, i.e., the transcriptional startsite referred to above. *Cis* elements occur within, near to, adjacent to, or at a distance from a particular promoter, but remain linked to the promoter sequence along the sequence of phosphodiester bonds which comprise the nucleotide sequence within which the promoter resides. *Cis* elements are not limited to promoters, but may be imparted to RNA sequences derived from transcription from DNA sequences of the present invention, wherein such RNA *cis* elements are involved in post transcriptional regulation of gene expression. For example, elements which are known as inverted repeat sequences can assist in the formation of hairpin structures which prevent, inhibit, or otherwise modulate the translational efficiency of the RNA sequence, or which regulate the survival of the RNA sequence. Other elements may function to bind ribosomes or components which enhance or suppress translational efficiency. *Cis* elements can be identified using known *cis* elements as a target sequence or target motif in the BLAST. Promoters of the present invention include homologues of *cis* elements known to effect gene regulation that show homology with the nucleic acid molecules of the present invention.

The isolated nucleic acid molecules of the present invention also include nucleic acid sequences that encode ribosomal RNA (rRNA), transfer RNA (tRNA) molecules, or other nucleic acid molecules which function to regulate gene expression, transcription, translation by acting alone or in combination with other cellular components in activating, inhibiting, terminating or anti-terminating gene expression functions, or by acting alone or in combination with other structural molecules to form or assist in the formation of said structural molecules.

It is contemplated by the inventors that the isolated nucleic acid molecules of the present invention also include those comprising a substantial portion of a nucleotide sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283 or complements thereof.

A "substantial portion" of a nucleotide sequence comprises enough of the sequence to afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. In general, gene specific oligonucleotide probes comprising 20-30 contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., in situ hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12-15 bases may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

It is also contemplated by the inventors that the isolated nucleic acid molecules of the present invention also include known types of modifications, for example, labels which are known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog. Other known modifications include internucleotide modifications, for example, those with uncharged linkages (methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.) and with charged linkages (phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, proteins (including nucleases, metabolic toxins, antibodies, signal peptides, poly-L-lysine, etc.), those with intercalators (acridine, psoralen, etc.), those containing chelators (metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, and those with modified linkages.

The nucleic acids of the present invention may be used to isolate nucleic acids encoding homologous proteins from the same or other species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA amplification as

exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase chain reaction, ligase chain reaction).

For example, genes encoding homologous proteins, either as cDNAs or genomic DNAs, could be isolated directly by using all or a portion of the nucleic acids of the present invention as DNA

5 hybridization probes to screen cDNA or genomic libraries from any desired organism employing methodology well known to those skilled in the art. Methods for forming such libraries are well known in the art. Specific oligonucleotide probes based upon the nucleic acids of the present invention can be designed and synthesized by methods known in the art. Moreover, the entire sequences of the nucleic acids can be used directly to synthesize DNA probes by methods known to the skilled artisan such as  
10 random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available in vitro transcription systems. In addition, specific primers can be designed and used to amplify a part or all of the sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full length cDNA or genomic DNAs under conditions of appropriate stringency.

15 Alternatively, the nucleic acids of interest can be amplified from nucleic acid samples using amplification techniques. For instance, the disclosed nucleic acids may be used to define a pair of primers that can be used with the polymerase chain reaction (Mullis, *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 51:263-273 (1986); Erlich *et al.*, EP 50,424; EP 84,796, EP 258,017, EP 237,362; Mullis, EP 201,184; Mullis *et al.*, US 4,683,202; Erlich, US 4,582,788; and Saiki, R. *et al.*, US 4,683,194) to amplify  
20 and obtain any desired nucleic acid or fragment directly from mRNA, from cDNA, from genomic libraries or cDNA libraries. PCR and other in vitro amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes.

25 In addition, two short segments of the nucleic acids of the present invention may be used in polymerase chain reaction protocols, for example, the RACE protocol (Frohman *et al.*, *Proc. Natl. Acad. Sci. USA* 85:8998 (1988)), to amplify longer nucleic acids encoding homologous genes from DNA or RNA from other sources.

30 Nucleic acids of interest may also be synthesized, either completely or in part, especially where it is desirable to provide plant-preferred sequences, by well-known techniques as described in the technical literature. See, e.g., Carruthers *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 47:411-418 (1982), and Adams *et al.*, *J. Am. Chem. Soc.* 105:661 (1983). Thus, all or a portion of the nucleic acids of the present invention may be synthesized using codons preferred by a selected plant host. Plant-preferred codons

may be determined, for example, from the codons used most frequently in the proteins expressed in a particular plant host species. Other modifications of the gene sequences may result in mutants having slightly altered activity.

Availability of the nucleotide sequences encoding *Bacillus thuringiensis* proteins facilitates immunological screening of DNA expression libraries. Synthetic polypeptides representing portions of the amino acid sequences of *Bacillus thuringiensis* proteins may be synthesized. These polypeptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for polypeptides or proteins comprising the amino acid sequences. These antibodies can be then be used to screen expression libraries to isolate genes of interest (Lemer, Adv. ImmunoL 36: 1 (1984); Sambrook et al., Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, (1989)). It is understood that people skilled in the art are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of antibodies (see, for example, Harlow and Lane, In *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1988)).

Another aspect of the present invention relates to a method for obtaining a nucleic acid comprising a nucleotide sequence encoding a *Bacillus thuringiensis* protein homologue listed in Table 1. In a preferred embodiment, the method of the present invention for obtaining a nucleic acid encoding all or a substantial portion of the amino acid sequence of a *Bacillus thuringiensis* protein homologue comprising: (a) probing an expression library with a hybridization probe comprising a nucleotide sequence encoding a polypeptide listed in Table 1; (b) identifying a DNA clone that hybridizes to the hybridization probe; (c) isolating the DNA clone identified in step (b); and (d) sequencing the DNA fragment that comprises the clone isolated in step (c) wherein the sequenced nucleic acid molecule encodes all or a substantial portion of the amino acid sequence of the *Bacillus thuringiensis* protein homologue.

In another preferred embodiment, the method of the present invention for obtaining a nucleic acid molecule encoding a substantial portion of an amino acid sequence of a *Bacillus thuringiensis* protein homologue comprising: (a) synthesizing a first and a second oligonucleotide primers corresponding to a portion of the coding sequence of a second nucleic acid molecule set forth in SEQ ID NO: 1 through SEQ ID NO: 8283; and (b) amplifying a DNA insert present in a cloning vector using the first and second oligonucleotide primers of step (a) wherein the amplified nucleic acid molecule encodes all or a substantial portion of the amino acid sequence of the *Bacillus thuringiensis* protein homologue.



### **Protein and Polypeptide Molecules**

The present invention, in another aspect, provides a substantially purified protein or polypeptide molecule comprising an amino acid sequence, wherein the amino acid sequence is defined as follows: (1) the amino acid sequence is encoded by a first nucleotide sequence which specifically hybridizes to the complement of a second nucleotide sequence set forth in SEQ ID NO: 1 through SEQ ID NO: 8283; or (2) the amino acid sequence is encoded by a third nucleotide sequence that is at least 50% identical to all or a substantial portion of a coding sequence located within SEQ ID NO: 1 through SEQ ID NO: 8283. In alternative embodiments, the third nucleotide sequence is at least 55% identical, at least 60% identical, at least 65% identical, at least 70% identical, at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical to all or a substantial portion of a coding sequence located within SEQ ID NO: 1 through SEQ ID NO: 8283. In a preferred embodiment, the third nucleotide sequence is 100% identical to all or a substantial portion of a coding sequence located within SEQ ID NO: 1 through SEQ ID NO: 8283.

The term "substantially purified protein or polypeptide molecule" refers to a protein or polypeptide molecule separated from substantially all other molecules normally associated with it in its native state. More preferably a substantially purified protein or polypeptide molecule is the predominant species present in a preparation. A substantially purified molecule may be greater than 60% free, preferably 75% free, more preferably 90% free, and most preferably 95% free from the other molecules (exclusive of solvent) present in the natural mixture.

It is well known in the art that proteins or polypeptides may undergo modification, including post-translational modifications, such as, but not limited to, disulfide bond formation, glycosylation, phosphorylation, or oligomerization. Thus, as used herein, the term "protein molecule" or "polypeptide molecule" includes any protein molecule that is modified by any biological or non-biological process. The terms "amino acid" and "amino acids" refer to all naturally occurring L-amino acids. This definition is meant to include norleucine, ornithine, homocysteine, and homoserine.

The polypeptides or proteins of the present invention may be produced via chemical synthesis, or more preferably, by expression in a suitable bacterial or eukaryotic host. Suitable methods for expression are described by Sambrook, *et al.*, (In: *Molecular Cloning, A Laboratory Manual, 2nd Edition*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989)), or similar texts.

The polypeptides or protein molecule of the present invention may also include fusion protein or polypeptide molecules. A protein or polypeptide molecule that comprises one or more additional polypeptide regions not derived from that protein molecule is a "fusion" protein or polypeptide molecule. Such molecules may be derivatized to contain carbohydrate or other moieties (such as keyhole limpet

hemocyanin, etc.). Fusion protein or polypeptide molecules of the present invention are preferably produced via recombinant means.

The protein or polypeptide molecules of the present invention may also include protein or polypeptide molecules encoded by all or a substantial portion of protein-encoding sequences in SEQ ID NO: 1 through SEQ ID NO: 8283 or complements thereof or, fragments or fusions thereof in which conservative, non-essential, or not relevant, amino acid residues have been added, replaced, or deleted. An example of such a homologue is the homologue protein from different strains or species. Such a homologue can be obtained by any of a variety of methods. For example, as indicated above, one or more of the disclosed sequences (all or a substantial portion of the protein-encoding sequences in SEQ ID NO: 1 through SEQ ID NO: 8283 or complements thereof) will be used to define a pair of primers that may be used to isolate the homologue-encoding nucleic acid molecules from any desired species. Such molecules can be expressed to yield homologues by recombinant means.

#### Antibodies

Another aspect of the present invention concerns antibodies, single-chain antigen binding molecules, or other proteins that specifically bind to one or more of the protein or polypeptide molecules of the present invention and their homologues, fusions or fragments. Such antibodies may be used to quantitatively or qualitatively detect the protein or polypeptide molecules of the present invention. As used herein, an antibody or polypeptide is said to "specifically bind" to a protein or polypeptide molecule of the present invention if such binding is not competitively inhibited by the presence of non-related molecules. In a preferred embodiment the antibodies of the present invention bind to protein or polypeptide molecules of the present invention, in a more preferred embodiment of the antibodies of the present invention bind to protein or polypeptide molecules derived from *Bacillus thuringiensis*.

Nucleic acid molecules that encode all or part of the protein or polypeptide of the present invention can be expressed, via recombinant means, to yield protein or polypeptides that can in turn be used to elicit antibodies that are capable of binding the expressed protein or polypeptide. Such antibodies may be used in immunoassays for that protein or polypeptide. Such protein or polypeptide-encoding molecules, or their fragments may be "fusion" molecules (*i.e.*, a part of a larger nucleic acid molecule) such that, upon expression, a fusion protein is produced. It may be desirable to derivatize the obtained antibodies, for example with a ligand group (such as biotin) or a detectable marker group (such as a fluorescent group, a radioisotope or an enzyme). Such antibodies may be used in immunoassays for that protein. In a preferred embodiment, such antibodies can be used to screen DNA expression libraries to isolate clones containing full-length insert of genes (Lemer, Adv. Immunol. 36: 1 (1984); Sambrook et

al., *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, (1989)).

The antibodies that specifically bind proteins and protein fragments of the present invention may be polyclonal or monoclonal, and may comprise intact immunoglobulins, or antigen binding portions of immunoglobulins (such as  $F(ab')$ ,  $F(ab')_2$  fragments), or single-chain immunoglobulins producible, for example, via recombinant means). It is understood that practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of antibodies (see, for example, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1988)).

In a preferred embodiment, the antibodies of the present invention specifically bind to one or more of the insect inhibitory polypeptides or proteins of the present invention. Such antibodies may be used to detect the presence of such insect inhibitory polypeptides or proteins in a sample.

The present invention also provide a method for detecting an insect inhibitory polypeptide or protein in a biological sample, the method generally comprising: (1) obtaining a biological sample; (2) contacting the sample with an antibody that specifically binds to the polypeptide or protein, under conditions effective to allow the formation of complexes; and (3) detecting the complexes so formed.

#### **Microbial Constructs and Transformed microbial Cells**

The nucleotide sequences of the present invention may be introduced into a wide variety of prokaryotic and eukaryotic microorganism hosts to express the *B. thuringiensis* polypeptide or protein of interest, particularly the insect inhibitory polypeptides or proteins of the present invention. The term "microorganism" includes prokaryotic and eukaryotic microbial species such as bacteria and fungi. Fungi include yeast and filamentous fungi. Illustrative prokaryotes, both Gram-negative and Gram-positive, include *Enterobacteriaceae*, such as *Escherichia*, *Erwinia*, *Shigella*, *Salmonella*, and *Proteus*; *Bacillaceae*; *Rhizobiceae*, such as *Rhizobium*; *Spirillaceae*, such as photobacterium, *Zymomonas*, *Serratia*, *Aeromonas*, *Vibrio*, *Desulfovibrio*, *Spirillum*; *Lactobacillaceae*; *Pseudomonadaceae*, such as *Pseudomonas* and *Acetobacter*; *Azotobacteraceae*, *Actinomycetales*, and *Nitrobacteraceae*. Among eukaryotes are fungi, such as *Phycomycetes* and *Ascomycetes*, which includes yeast, such as *Saccharomyces* and *Schizosaccharomyces*; and *Basidiomycetes* yeast, such as *Rhodotorula*, *Aureobasidium*, *Sporobolomyces*, and the like.

For the purpose of plant protection against insects, a large number of microorganisms known to inhabit the phylloplane (the surface of the plant leaves) and/or the rhizosphere (the soil surrounding plant

roots) of a wide variety of important crops may also be desirable host cells for manipulation, propagation, storage, delivery and/or mutagenesis of the disclosed recombinant constructs. These microorganisms include bacteria, algae, and fungi. Of particular interest are microorganisms, such as bacteria, *e.g.*, genera *Bacillus* (including the species and subspecies *B. thuringiensis kurstaki* HD-1, *B. thuringiensis kurstaki* HD-73, *B. thuringiensis sotto*, *B. thuringiensis berliner*, *B. thuringiensis thuringiensis*, *B. thuringiensis tolworthi*, *B. thuringiensis dendrolimus*, *B. thuringiensis alesti*, *B. thuringiensis galleriae*, *B. thuringiensis aizawai*, *B. thuringiensis subtoxicus*, *B. thuringiensis entomocidus*, *B. thuringiensis tenebrionis* and *B. thuringiensis san diego*); *Pseudomonas*, *Erwinia*, *Serratia*, *Klebsiella*, *Zanthomonas*, *Streptomyces*, *Rhizobium*, *Rhodopseudomonas*, *Methylophilus*, *Agrobacterium*, *Acetobacter*, *Lactobacillus*, *Arthrobacter*, *Azotobacter*, *Leuconostoc*, and *Alcaligenes*; fungi, particularly yeast, *e.g.*, genera *Saccharomyces*, *Cryptococcus*, *Kluyveromyces*, *Sporobolomyces*, *Rhodotorula*, and *Aureobasidium*. Of particular interest are such phytosphere bacterial species as *Pseudomonas syringae*, *Pseudomonas fluorescens*, *Serratia marcescens*, *Acetobacter xylinum*, *Agrobacterium tumefaciens*, *Rhodobacter sphaeroides*, *Xanthomonas campestris*, *Rhizobium melioli*, *Alcaligenes eutrophus*, and *Azotobacter vinlandii*; and phytosphere yeast species such as *Rhodotorula rubra*, *R. glutinis*, *R. marina*, *R. aurantiaca*, *Cryptococcus albidus*, *C. diffluens*, *C. laurentii*, *Saccharomyces rosei*, *S. pretoriensis*, *S. cerevisiae*, *Sporobolomyces roseus*, *S. odoratus*, *Kluyveromyces veronae*, and *Aureobasidium pollulans*.

It is well known that exogenous nucleic acids encoding polypeptides of interest can be introduced into a microbial host cell, such as a bacterial cell or a fungal cell, using a recombinant construct. The present invention also relates to a fungal or bacterial recombinant construct comprising a structural nucleotide sequence encoding a *B. thuringiensis* protein or polypeptide. In a preferred embodiment, the structural nucleotide sequence encodes an insect inhibitory protein or polypeptide of the present invention. The present invention also relates to a bacterial or fungal cell comprising a bacterial or fungal recombinant vector. The present invention also relates to methods for obtaining a recombinant bacterial or fungal host cell, comprising introducing into a bacterial or fungal host cell an exogenous nucleic acid molecule.

The bacterial recombinant vector may be any vector which can be conveniently subjected to recombinant DNA procedures. The choice of a vector will typically depend on the compatibility of the vector with the bacterial host cell into which the vector is to be introduced. The vector may be a linear or a closed circular plasmid. The bacterial recombinant vector may be a linear or a closed circular plasmid. The vector system may be a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the bacterial host. In addition, the bacterial vector may be an expression vector. Nucleic acid molecules encoding *B. thuringiensis* proteins or

polypeptide can, for example, be suitably inserted into a replicable vector for expression in a bacterium under the control of a suitable promoter for that bacterium. Many vectors are available for this purpose, and selection of the appropriate vector will depend mainly on the size of the nucleic acid to be inserted into the vector and the particular host cell to be transformed with the vector. Each vector contains various components depending on its function (amplification of DNA or expression of DNA) and the particular host cell with which it is compatible. The vector components for bacterial transformation generally include, but are not limited to, one or more of the following: a signal sequence, an origin of replication, one or more selectable marker genes, and an inducible promoter allowing the expression of exogenous DNA.

In general, plasmid vectors containing replicon and control sequences that are derived from species compatible with the host cell are used in connection with bacterial hosts. The vector ordinarily carries a replication site, as well as marking sequences that are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using pBR322, a plasmid derived from an *E. coli* species (see, e.g., Bolivar *et al.*, *Gene* 2:95 (1977)). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage, also generally contains, or is modified to contain, promoters that can be used by the microbial organism for expression of the selectable marker genes.

Nucleic acid molecules encoding *B. thuringiensis* proteins or polypeptides may be expressed not only directly, but also as a fusion with another polypeptide, preferably a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the polypeptide encoding DNA that is inserted into the vector. The heterologous signal sequence selected should be one that is recognized and processed (i.e., cleaved by a signal peptidase) by the host cell. For bacterial host cells that do not recognize and process the native polypeptide signal sequence, the signal sequence is substituted by a bacterial signal sequence selected, for example, from the group consisting of the alkaline phosphatase,  $\beta$ -lactamase, or heat-stable enterotoxin II leaders and the like.

Both expression and cloning vectors contain a nucleotide sequence that enables the vector to replicate in one or more selected host cells. Generally, in cloning vectors this sequence is one that enables the vector to replicate independently of the host chromosomal DNA, and includes origins of replication or autonomously replicating sequences. Such sequences are well known for a variety of bacteria.

Expression and cloning vectors also generally contain a selection gene, also termed a selectable marker. This gene encodes a protein necessary for the survival or growth of transformed host cells grown

in a selective culture medium. Host cells not transformed with the vector containing the selection gene will not survive in the culture medium. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*. One example of a selection scheme utilizes a drug to arrest growth of a host cell. Those cells that are successfully transformed with a heterologous protein homologue or fragment thereof produce a protein conferring drug resistance and thus survive the selection regimen.

The expression vector for producing a polypeptide can also contains an inducible promoter that is recognized by the host bacterial organism and is operably linked to the nucleic acid encoding, for example, the nucleic acid molecule encoding the *B. thuringiensis* protein or polypeptide of interest. Inducible promoters suitable for use with bacterial hosts include the  $\beta$ -lactamase, *E. coli*  $\lambda$  phage P<sub>L</sub> and P<sub>R</sub>, and *E. coli* galactose, arabinose, alkaline phosphatase, tryptophan (trp), and lactose operon promoter systems and variations thereof (Chang *et al.*, *Nature* 275:615 (1978); Goeddel *et al.*, *Nature* 281:544 (1979); Guzman *et al.*, *J. Bacteriol.* 174:7716-7728 (1992); Goeddel, *Nucleic Acids Res.* 8:4057 (1980); EP 36,776) and hybrid promoters such as the tac promoter (deBoer *et al.*, *Proc. Natl. Acad. Sci. (USA)* 80:21-25 (1983)). However, other known bacterial inducible promoters are suitable (Siebenlist *et al.*, *Cell* 20:269 (1980)).

The term "operably linked", as used in reference to a regulatory sequence and a structural nucleotide sequence, means that the regulatory sequence causes regulated expression of the operably linked structural nucleotide sequence. "Expression" refers to the transcription and stable accumulation of sense or antisense RNA derived from the nucleic acid of the present invention. Expression may also refer to translation of mRNA into a polypeptide or protein. "Sense" RNA refers to RNA transcript that includes the mRNA and so can be translated into protein by the cell. "Antisense RNA" refers to a RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene (U.S. Pat. No. 5,107,065, incorporated herein by reference). The complementarity of an antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-translated sequence, introns, or the coding sequence. "RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from post-transcriptional processing of the primary transcript and is referred to as the mature RNA.

"Regulatory sequences" or "control elements" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-translated sequences) of a structural nucleotide sequence, and which influence the transcription, RNA processing or stability, or translation of the associated structural nucleotide sequence. Regulatory sequences may include promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

The promoter sequence may consist of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a DNA sequence which can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions.

Promoters for use in bacterial systems also generally contain a Shine-Dalgarno (S.D.) sequence or a consensus sequence thereof operably linked to the DNA encoding the polypeptide of interest. The promoter can be removed from the bacterial source DNA by restriction enzyme digestion and inserted into the vector containing the desired DNA coding sequence, or vice versa.

Alternatively, the expression constructs can be integrated into the bacterial genome with an integrating vector. Integrating vectors typically contain at least one sequence homologous to the bacterial chromosome that allows the vector to integrate. Integrations appear to result from recombinations between homologous DNA in the vector and the bacterial chromosome. For example, integrating vectors constructed with DNA from various *Bacillus* strains integrate into the *Bacillus* chromosome (E.P.O. Pub. No. 127,328). Integrating vectors may also be comprised of bacteriophage or transposon sequences.

Construction of suitable vectors containing one or more of the above-listed components employs standard recombinant DNA techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and re-ligated in the form desired to generate the plasmids required. Examples of available bacterial expression vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as Bluescript<sup>TM</sup> (Stratagene, La Jolla, CA), in which, for example, a *B. thuringiensis* protein or polypeptide of the present invention, may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of  $\beta$ -galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke and Schuster *J. Biol. Chem.* 264:5503-5509 (1989)); and the like. pGEX vectors (Promega, Madison Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily

be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems are designed to include heparin, thrombin or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

5 It is, of course, necessary to select the appropriate bacteria taking into consideration replicability of the replicon in the cells of a bacterium. For example, *E. coli*, *Serratia*, or *Salmonella* species can be suitably used as the host when well known plasmids such as pBR322, pBR325, pACYC177, or pKN410 are used to supply the replicon. *E. coli* strain W3110 is a preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell should secrete  
10 minimal amounts of proteolytic enzymes.

Expression vectors for *B. thuringiensis* and related bacilli contain an origin of DNA replication and associated replication factors and a selectable marker gene that will be expressed in *Bacillus* and allow selection of transformed cells. These *Bacillus*-specific plasmid functions may also be present on an *E. coli* or other broad host range cloning vector, forming a plasmid that can replicate and be selected in  
15 either *E. coli* or *B. thuringiensis*, and thus comprising a shuttle vector. Such vectors are well known in the art and usually also contain a *B. thuringiensis*  $\delta^E$  promoter adjacent to one or more convenient cloning sites to allow cloning and expression of exogenous genes during sporulation. Such a promoter is typically found associated with many  $\delta$ -endotoxin genes.

Host cells are transfected and preferably transformed with the above-described vectors and  
20 cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

Numerous methods of transfection are known to the ordinarily skilled artisan, for example, calcium phosphate and electroporation. Depending on the host cell used, transformation is done using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as  
25 described in section 1.82 of Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Laboratory Press, (1989), is generally used for bacterial cells that contain substantial cell-wall barriers. Another method for transformation employs polyethylene glycol/DMSO, as described in Chung and Miller (Chung and Miller, *Nucleic Acids Res.* 16:3580 (1988)). Yet another method is the use of the technique termed electroporation. In addition, bacterial cells can be readily transformed using  
30 various forms of phages (i.e., transducing, temperate, lytic and lysogenic), suicide vectors for inserting DNA directly into the chromosome, and through homologous recombination using either phages, suicide vectors or linear DNA.



Bacterial cells used to produce the polypeptide of interest for purposes of this invention are cultured in suitable media in which the promoters for the nucleic acid encoding the heterologous polypeptide can be artificially induced as described generally, e.g., in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Laboratory Press, (1989). Examples of  
 5 suitable media are given in U.S. Pat. Nos. 5,304,472 and 5,342,763.

A yeast recombinant construct can typically include one or more of the following: a promoter sequence, fusion partner sequence, leader sequence, transcription termination sequence, a selectable marker. These elements can be combined into an expression cassette, which may be maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host,  
 10 such as yeast or bacteria. The replicon may have two replication systems, thus allowing it to be maintained, for example, in yeast for expression and in a procaryotic host for cloning and amplification. Examples of such yeast-bacteria shuttle vectors include YEp24 (Botstein *et al.*, *Gene*, 8:17-24 (1979)), pCI/1 (Brake *et al.*, *Proc. Natl. Acad. Sci USA*, 81:4642-4646 (1984)), and YRp17 (Stinchcomb *et al.*, *J. Mol. Biol.*, 158:157 (1982)). In addition, a replicon may be either a high or low copy number plasmid. A  
 15 high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and typically about 10 to about 150. A host containing a high copy number plasmid will preferably have at least about 10, and more preferably at least about 20.

Useful yeast promoter sequences can be derived from genes encoding enzymes in the metabolic pathway. Examples of such genes include alcohol dehydrogenase (ADH) (E.P.O. Pub. No. 284044),  
 20 enolase, glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-dehydrogenase (GAP or GAPDH), hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, and pyruvate kinase (PyK) (E.P.O. Pub. No. 329203). The yeast PHO5 gene, encoding acid phosphatase, also provides useful promoter sequences (Myanohara *et al.*, *Proc. Natl. Acad. Sci. USA*, 80:1 (1983)). In addition, synthetic promoters which do not occur in nature also function as yeast promoters. Examples of such hybrid  
 25 promoters include the ADH regulatory sequence linked to the GAP transcription activation region (U.S. Pat. Nos. 4,876,197 and 4,880,734). Other examples of hybrid promoters include promoters which consist of the regulatory sequences of either the ADH2, GAL4, GAL10, or PHO5 genes, combined with the transcriptional activation region of a glycolytic enzyme gene such as GAP or PyK (E.P.O. Pub. No. 164556). Furthermore, a yeast promoter can include naturally occurring promoters of non-yeast origin  
 30 that have the ability to bind yeast RNA polymerase and initiate transcription. Examples of such promoters include, inter alia, (Cohen *et al.*, *Proc. Natl. Acad. Sci. USA*, 77:1078 (1980); Henikoff *et al.*, *Nature* 283:835 (1981); Hollenberg *et al.*, *Curr. Topics Microbiol. Immunol.*, 96:119 (1981); Mercerau-Puigalon *et al.*, *Gene*, 11:163 (1980); and Panthier *et al.*, *Curr. Genet.*, 2:109 (1980)).

Intracellularly expressed fusion proteins provide an alternative to direct expression of the polypeptides of interest. Typically, a DNA sequence encoding the N-terminal portion of a stable protein, a fusion partner, is fused to the 5' end of heterologous structural nucleotide sequence encoding the desired polypeptide. Upon expression, this construct will provide a fusion of the two amino acid sequences. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. See, e.g., E.P.O. Pub. No. 196056. Another example is a ubiquitin fusion protein. Such a ubiquitin fusion protein preferably retains a site for a processing enzyme (e.g. ubiquitin-specific processing protease) to cleave the ubiquitin from the polypeptide of the present invention. Through this method, therefore, a mature polypeptide can be isolated [see, P.C.T. WO 88/024066].

Alternatively, polypeptides or proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion in yeast of the polypeptides. Preferably, there are processing sites encoded between the leader fragment and the polypeptide-encoding sequence fragment that can be cleaved either in vivo or in vitro. The leader sequence fragment typically encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell.

DNA encoding suitable signal sequences can be derived from genes for secreted yeast proteins, such as the yeast invertase gene (E.P.O. Pub. No. 12873; J.P.O. Pub. No. 62,096,086) and the A-factor gene (U.S. Pat. No. 4,588,684). Alternatively, leaders of non-yeast origin, such as an interferon leader, exist that also provide for secretion in yeast (E.P.O. Pub. No. 60057).

A preferred class of secretion leaders are those that employ a fragment of the yeast alpha-factor gene, which contains both a "pre" signal sequence, and a "pro" region. The types of alpha-factor fragments that can be employed include the full-length pre-pro alpha factor leader (about 83 amino acid residues) as well as truncated alpha-factor leaders (typically about 25 to about 50 amino acid residues) (U.S. Pat. Nos. 4,546,083 and 4,870,008; and E.P.O. Pub. No. 324274). Additional leaders employing an alpha-factor leader fragment that provides for secretion include hybrid alpha-factor leaders made with a pre-sequence of a first yeast, but a pro-region from a second yeast alpha factor. See, e.g., P.C.T. WO 89/02463.

Examples of transcription terminator sequence and other yeast-recognized termination sequences, such as those coding for glycolytic enzymes, are known to those of skill in the art.

Alternatively, the expression constructs can be integrated into the yeast genome with an integrating vector. Integrating vectors typically contain at least one sequence homologous to a yeast chromosome that allows the vector to integrate, and preferably contain two homologous sequences flanking the expression construct. Integrations appear to result from recombinations between homologous

DNA in the vector and the yeast chromosome (Orr-Weaver et al., Methods in Enzymol., 101:228-245 (1983)). An integrating vector may be directed to a specific locus in yeast by selecting the appropriate homologous sequence for inclusion in the vector. See Orr-Weaver et al., supra. One or more expression constructs may integrate, possibly affecting levels of recombinant protein produced (Rine et al., Proc. Natl. Acad. Sci. USA, 80:6750 (1983)). The chromosomal sequences included in the vector can occur either as a single segment in the vector, which results in the integration of the entire vector, or as two segments homologous to adjacent segments in the chromosome and flanking the expression construct in the vector, which results in the stable integration of only the expression construct.

Expression and transformation vectors, either extrachromosomal replicons or integrating vectors, have been developed for transformation into many yeasts. For example, expression vectors have been developed for, inter alia, the following yeasts: *Candida albicans* (Kurtz, et al., Mol. Cell. Biol., 6:142 (1986)), *Candida maltosa* (Kunze et al., J. Basic Microbiol., 25:141 (1985)); *Hansenula polymorpha* (Gleeson et al., J. Gen. Microbiol. 132:3459 (1986); Roggenkamp et al., Mol. Gen. Genet. 202:302 (1986)); *Kluyveromyces fragilis* (Das et al., J. Bacteriol. 158:1165 (1984)); *Kluyveromyces lactis* (De Louvencourt et al., J. Bacteriol. 154:737 (1983); Van den Berg et al., Bio/Technology 8:135 (1990)); *Pichia guilliermondii* (Kunze et al., J. Basic Microbiol. 25:141 (1985)); *Pichia pastoris* (Cregg et al., Mol. Cell. Biol. 5:3376 (1985); U.S. Pat. Nos. 4,837,148 and 4,929,555); *Saccharomyces cerevisiae* (Hinnen et al., Proc. Natl. Acad. Sci. USA 75:1929 (1978); Ito et al., J. Bacteriol. 153:163 (1983)); *Schizosaccharomyces pombe* (Beach and Nurse, Nature 300:706 (1981)); and *Yarrowia lipolytica* (Davidow, et al., Curr. Genet. 10:380471 (1985); and Gaillardin et al., Curr. Genet. 10:49 (1985)).

Methods of introducing exogenous nucleic acids into yeast hosts are well-known in the art, and typically include either the transformation of spheroplasts or of intact yeast cells treated with alkali cations. Transformation procedures usually vary with the yeast species to be transformed. See e.g., Kurtz et al., Mol. Cell. Biol. 6:142 (1986); Kunze et al., J. Basic Microbiol. 25:141 (1985) for *Candida*. See, e.g., Gleeson et al., J. Gen. Microbiol. 132:3459 (1986); Roggenkamp et al., Mol. Gen. Genet. 202:302 (1986) for *Hansenula*. See, e.g., Das et al., J. Bacteriol. 158:1165 (1984); De Louvencourt et al., J. Bacteriol. 154:1165 (1983); Van den Berg et al., Bio/Technology 8:135 (1990) for *Kluyveromyces*. See, e.g., Cregg et al., Mol. Cell. Biol. 5:3376 (1985); Kunze et al., J. Basic Microbiol. 25:141 (1985); U.S. Pat. Nos. 4,837,148 and 4,929,555 for *Pichia*. See, e.g., Hinnen et al., Proc. Natl. Acad. Sci. USA 75:1929 (1978); Ito et al., J. Bacteriol. 153:163 (1983) for *Saccharomyces*. See, e.g., Beach and Nurse, Nature 300:706 (1981) for *Schizosaccharomyces*. See, e.g., Davidow et al., Curr. Genet. 10:39 (1985); Gaillardin et al., Curr. Genet. 10:49 (1985) for *Yarrowia*.

In order to obtain expression polypeptides or proteins of interest, recombinant microbial host cells derived from the transformants are incubated under conditions which allow expression of the recombinant polypeptide-encoding sequence. These conditions will vary, dependent upon the host cell selected. However, the conditions are readily ascertainable to those of ordinary skill and knowledge in the art.

Detection of polypeptides expressed in the transformed host cell may be performed by several methods. For example, a polypeptide or protein may be detected by its immunological reactivity with antibodies.

Polypeptides or proteins of the present invention may be isolated from the cell by lysis, if formed intracellularly, or isolated from the culture medium, if secreted, by conventional methods.

### **Plant Constructs and Plant Transformants**

Nucleic acid molecules of the present invention may be used in plant transformation or transfection. Exogenous genetic material may be transferred into a plant cell and the plant cell regenerated into a whole, fertile or sterile plant. Such genetic material may be transferred into either monocotyledons and dicotyledons including but not limited to the plants, alfalfa, *Arabidopsis*, barley, *Brassica*, broccoli, cabbage, citrus, cotton, garlic, oat, oilseed rape, onion, canola, flax, maize, an ornamental plant, pea, peanut, pepper, potato, rice, rye, sorghum, soybean, strawberry, sugarcane, sugarbeet, tomato, wheat, poplar, pine, fir, eucalyptus, apple, lettuce, lentils, grape, banana, tea, turf grasses, sunflower, oil palm, *Phaseolus* etc. Particularly preferred plants would include, *Arabidopsis*, barley, cotton, oat, oilseed rape, rice, maize, soybean, canola, ornamentals, sugarcane, sugarbeet, tomato, potato, wheat and turf grasses (*See specifically*, Chistou, *Particle Bombardment for Genetic Engineering of Plants*, Biotechnology Intelligence Unit, Academic Press, San Diego, CA (1996) and generally Chistou, *Particle Bombardment for Genetic Engineering of Plants*, Biotechnology Intelligence Unit, Academic Press, San Diego, CA (1996)).

The present invention also relates to a transformed plant cell or plant comprising in its genome an exogenous nucleic acid encoding one or more *B. thuringiensis* proteins or polypeptides of the present invention. The present invention also relates to methods for creating a transgenic plant in which one or more *B. thuringiensis* proteins or polypeptides of the present invention are overexpressed.

As used herein, the term exogenous genetic material means any genetic material, whether naturally occurring or otherwise, from any source that is capable of being inserted into any organism. In a preferred embodiment the exogenous genetic material includes an isolated nucleic acid molecule or fragment thereof of the present invention. In another preferred embodiment of the present invention

exogenous genetic material includes a regulatory element, promoter or partial promoter of the present invention. In a more preferred embodiment of the present invention exogenous genetic material includes a regulatory element, promoter or partial promoter of the present invention and a nucleic acid molecule encoding a *B. thuringiensis* protein or fragments thereof. In a further more preferred embodiment of the present invention exogenous genetic material includes a regulatory element, promoter or partial promoter of the present invention and a nucleic acid molecule of the present invention having a sequence encoding a polypeptide or protein or fragment thereof set forth in Table 1.

One skilled in the art understands that an exogenous nucleic acid can be a heterologous nucleic acid derived from a different species than the species into which the nucleic acid is introduced or can be a nucleic acid derived from the same species as the species into which it is introduced.

The term "overexpression" refers to the expression of a polypeptide or protein encoded by an exogenous nucleic acid introduced into a host cell, wherein said polypeptide or protein is either not normally present in the host cell, or wherein said polypeptide or protein is present in said host cell at a higher level than that normally expressed from the endogenous gene encoding said polypeptide or protein. By "endogenous gene" refers to a native gene in its natural location in the genome of an organism.

The term "genome" as it applies to plant cells encompasses not only chromosomal DNA found within the nucleus, but organelle DNA found within subcellular components of the cell. DNAs of the present invention introduced into plant cells can therefore be either chromosomally integrated or organelle-localized. The term "genome" as it applies to bacteria encompasses both the chromosome and plasmids within a bacterial host cell. Encoding DNAs of the present invention introduced into bacterial host cells can therefore be either chromosomally integrated or plasmid-localized.

Transfer of a nucleic acid molecule that encodes for a protein can result in overexpression of that protein in a transformed cell or transgenic plant. One or more of the proteins or fragments thereof encoded by nucleic acid molecules of the present invention may be overexpressed in a transformed cell or transformed plant. Such overexpression may be the result of transient or stable transfer of the exogenous material. In a preferred embodiment of the present invention, one or more of the *B. thuringiensis* proteins or fragments thereof are overexpressed in a transformed cell or transgenic plant.

Exogenous genetic material may be transferred into a plant cell by the use of a DNA vector or construct designed for such a purpose. Vectors have been engineered for transformation of large DNA inserts into plant genomes. Binary bacterial artificial chromosomes have been designed to replicate in both *E. coli* and *A. tumefaciens* and have all of the features required for transferring large inserts of DNA into plant chromosomes Choi and Wing, <http://genome.clemson.edu/protocols2-nj.html>, July, 1998.

ApBACwch system has been developed to achieve site-directed integration of DNA into the genome. A 150 kb cotton BAC DNA is reported to have been transferred into a specific *lox* site in tobacco by biolistic bombardment and *Cre-lox* site specific recombination.

A construct or vector may also include a plant promoter to express the protein or protein fragment of choice. A number of promoters which are active in plant cells have been described in the literature. These include the nopaline synthase (NOS) promoter (Ebert *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 84:5745-5749 (1987)), the octopine synthase (OCS) promoter (which are carried on tumor-inducing plasmids of *Agrobacterium tumefaciens*), the caulimovirus promoters such as the cauliflower mosaic virus (CaMV) 19S promoter (Lawton *et al.*, *Plant Mol. Biol.* 9:315-324 (1987)) and the CaMV 35S promoter (Odell *et al.*, *Nature* 313:810-812 (1985)), the figwort mosaic virus 35S-promoter, the light-inducible promoter from the small subunit of ribulose-1,5-bis-phosphate carboxylase (ssRUBISCO), the Adh promoter (Walker *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 84:6624-6628 (1987)), the sucrose synthase promoter (Yang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 87:4144-4148 (1990)), the R gene complex promoter (Chandler *et al.*, *The Plant Cell* 1:1175-1183 (1989)), and the chlorophyll a/b binding protein gene promoter, etc. These promoters have been used to create DNA constructs which have been expressed in plants; *see, e.g.*, PCT publication WO 84/02913.

Promoters which are known or are found to cause transcription of DNA in plant cells can be used in the present invention. Such promoters may be obtained from a variety of sources such as plants and plant viruses. It is preferred that the particular promoter selected should be capable of causing sufficient expression to result in the production of an effective amount of protein to cause the desired phenotype. In addition to promoters which are known to cause transcription of DNA in plant cells, other promoters may be identified for use in the current invention by screening a plant cDNA library for genes which are selectively or preferably expressed in the target tissues or cells.

For the purpose of expression in source tissues of the plant, such as the leaf, seed, root or stem, it is preferred that the promoters utilized in the present invention have relatively high expression in these specific tissues. For this purpose, one may choose from a number of promoters for genes with tissue- or cell-specific or -enhanced expression. Examples of such promoters reported in the literature include the chloroplast glutamine synthetase GS2 promoter from pea (Edwards *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 87:3459-3463 (1990)), the chloroplast fructose-1,6-biphosphatase (FBPase) promoter from wheat (Lloyd *et al.*, *Mol. Gen. Genet.* 225:209-216 (1991)), the nuclear photosynthetic ST-LS1 promoter from potato (Stockhaus *et al.*, *EMBO J.* 8:2445-2451 (1989)), the phenylalanine ammonia-lyase (PAL) promoter and the chalcone synthase (CHS) promoter from *B. thuringiensis thaliana*. Also reported to be active in photosynthetically active tissues are the ribulose-1,5-bisphosphate carboxylase (RbcS) promoter from

eastern larch (*Larix laricina*), the promoter for the *cab* gene, *cab6*, from pine (Yamamoto *et al.*, *Plant Cell Physiol.* 35:773-778 (1994)), the promoter for the *Cab-1* gene from wheat (Fejes *et al.*, *Plant Mol. Biol.* 15:921-932 (1990)), the promoter for the *CAB-1* gene from spinach (Lubberstedt *et al.*, *Plant Physiol.* 104:97-1006 (1994)), the promoter for the *cab1R* gene from rice (Luan *et al.*, *Plant Cell.* 4:971-981 (1992)), the pyruvate, orthophosphate dikinase (PPDK) promoter from maize (Matsuoka *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 90:9586-9590 (1993)), the promoter for the tobacco *Lhcb1\*2* gene (Cerdan *et al.*, *Plant Mol. Biol.* 33:245-255. (1997)), the *A. nidulans thaliana* SUC2 sucrose-H<sup>+</sup> symporter promoter (Truernit *et al.*, *Planta.* 196:564-570 (1995)), and the *th spergillus thaliana* SUC2 sucrose-H<sup>+</sup> symporter promoter (Truernit *et al.*, *Planta.* 196:564-570 (1995)), and the promoter for the thylakoid membrane proteins from spinach (*psaD*, *psaF*, *psaE*, *PC*, *FNR*, *atpC*, *atpD*, *cab*, *rbcS*). Other promoters for the chlorophyll a/b-binding proteins may also be utilized in the present invention, such as the promoters for *LhcB* gene and *PsbP* gene from white mustard (*Sinapis alba*; Kretsch *et al.*, *Plant Mol. Biol.* 28:219-229 (1995)).

For the purpose of expression in sink tissues of the plant, such as the tuber of the potato plant, the fruit of tomato, or the seed of maize, wheat, rice, and barley, it is preferred that the promoters utilized in the present invention have relatively high expression in these specific tissues. A number of promoters for genes with tuber-specific or -enhanced expression are known, including the class I patatin promoter (Bevan *et al.*, *EMBO J.* 8:1899-1906 (1986); Jefferson *et al.*, *Plant Mol. Biol.* 14:995-1006 (1990)), the promoter for the potato tuber ADPGPP genes, both the large and small subunits, the sucrose synthase promoter (Salanoubat and Belliard, *Gene.* 60:47-56 (1987), Salanoubat and Belliard, *Gene.* 84:181-185 (1989)), the promoter for the major tuber proteins including the 22 kd protein complexes and proteinase inhibitors (Hannapel, *Plant Physiol.* 101:703-704 (1993)), the promoter for the granule bound starch synthase gene (GBSS) (Visser *et al.*, *Plant Mol. Biol.* 17:691-699 (1991)), and other class I and II patatins promoters (Koster-Topfer *et al.*, *Mol. Gen. Genet.* 219:390-396 (1989); Mignery *et al.*, *Gene.* 62:27-44 (1988)).

Other promoters can also be used to express a fructose 1,6 bisphosphate aldolase gene in specific tissues, such as seeds or fruits. The promoter for  $\beta$ -conglycinin (Chen *et al.*, *Dev. Genet.* 10:112-122 (1989)) or other seed-specific promoters such as the napin and phaseolin promoters, can be used. The zeins are a group of storage proteins found in maize endosperm. Genomic clones for zein genes have been isolated (Pedersen *et al.*, *Cell* 29:1015-1026 (1982)), and the promoters from these clones, including the 15 kD, 16 kD, 19 kD, 22 kD, 27 kD, and gamma genes, could also be used. Other promoters known to function, for example, in maize, include the promoters for the following genes: *waxy*, *Brittle*, *Shrunken 2*, Branching enzymes I and II, starch synthases, debranching enzymes, oleosins, glutelins, and sucrose

synthases. A particularly preferred promoter for maize endosperm expression is the promoter for the glutelin gene from rice, more particularly the Osgt-1 promoter (Zheng *et al.*, *Mol. Cell Biol.* 13:5829-5842 (1993)). Examples of promoters suitable for expression in wheat include those promoters for the ADPglucose pyrophosphorylase (ADPGPP) subunits, the granule bound and other starch synthases, the branching and debranching enzymes, the embryogenesis-abundant proteins, the gliadins, and the glutenins. Examples of such promoters in rice include those promoters for the ADPGPP subunits, the granule bound and other starch synthases, the branching enzymes, the debranching enzymes, sucrose synthases, and the glutelins. A particularly preferred promoter is the promoter for rice glutelin, Osgt-1. Examples of such promoters for barley include those for the ADPGPP subunits, the granule bound and other starch synthases, the branching enzymes, the debranching enzymes, sucrose synthases, the hordeins, the embryo globulins, and the aleurone specific proteins.

Root specific promoters may also be used. An example of such a promoter is the promoter for the acid chitinase gene (Samac *et al.*, *Plant Mol. Biol.* 25:587-596 (1994)). Expression in root tissue could also be accomplished by utilizing the root specific subdomains of the CaMV35S promoter that have been identified (Lam *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:7890-7894 (1989)). Other root cell specific promoters include those reported by Conkling *et al.* (Conkling *et al.*, *Plant Physiol.* 93:1203-1211 (1990)).

Additional promoters that may be utilized are described, for example, in U.S. Patent Nos. 5,378,619; 5,391,725; 5,428,147; 5,447,858; 5,608,144; 5,608,144; 5,614,399; 5,633,441; 5,633,435 and 4,633,436, all of which are herein incorporated in their entirety. In addition, a tissue specific enhancer may be used (Fromm *et al.*, *The Plant Cell* 1:977-984 (1989)). It is further understood that one or more of the promoters of the present invention may be used.

Constructs or vectors may also include, with the coding region of interest, a nucleic acid sequence that acts, in whole or in part, to terminate transcription of that region. For example, such sequences have been isolated including the Tr7 3' sequence and the nos 3' sequence (Ingelbrecht *et al.*, *The Plant Cell* 1:671-680 (1989); Bevan *et al.*, *Nucleic Acids Res.* 11:369-385 (1983)), or the like. It is understood that one or more sequences of the present invention that act, to terminate transcription may be used.

A vector or construct may also include other regulatory elements. Examples of such include the Adh intron 1 (Callis *et al.*, *Genes and Develop.* 1:1183-1200 (1987)), the sucrose synthase intron (Vasil *et al.*, *Plant Physiol.* 91:1575-1579 (1989)) and the TMV omega element (Gallie *et al.*, *The Plant Cell* 1:301-311 (1989)). These and other regulatory elements may be included when appropriate. It is also understood that one or more of the regulatory regions of the present invention may be used.



A vector or construct may also include a selectable marker. Selectable markers may also be used to select for plants or plant cells that contain the exogenous genetic material. Examples of such include, but are not limited to, a neo gene (Potrykus *et al.*, *Mol. Gen. Genet.* 199:183-188 (1985)) which codes for kanamycin resistance and can be selected for using kanamycin, G418, etc.; a bar gene which codes for bialaphos resistance; a mutant EPSP synthase gene (Hinchee *et al.*, *Bio/Technology* 6:915-922 (1988)) which encodes glyphosate resistance; a nitrilase gene which confers resistance to bromoxynil (Stalker *et al.*, *J. Biol. Chem.* 263:6310-6314 (1988)); a mutant acetolactate synthase gene (ALS) which confers imidazolinone or sulphonylurea resistance (European Patent Application 154,204 (Sept. 11, 1985)); and a methotrexate resistant DHFR gene (Thillet *et al.*, *J. Biol. Chem.* 263:12500-12508 (1988)).

A vector or construct may also include a transit peptide. Incorporation of a suitable chloroplast transit peptide may also be employed (European Patent Application Publication Number 0218571). Translational enhancers may also be incorporated as part of the vector DNA. DNA constructs could contain one or more 5' non-translated leader sequences which may serve to enhance expression of the gene products from the resulting mRNA transcripts. Such sequences may be derived from the promoter selected to express the gene or can be specifically modified to increase translation of the mRNA. Such regions may also be obtained from viral RNAs, from suitable eukaryotic genes, or from a synthetic gene sequence. For a review of optimizing expression of transgenes, see Koziel *et al.*, *Plant Mol. Biol.* 32:393-405 (1996).

A vector or construct may also include a screenable marker. Screenable markers may be used to monitor expression. Exemplary screenable markers include a  $\beta$ -glucuronidase or uidA gene (GUS) which encodes an enzyme for which various chromogenic substrates are known (Jefferson, *Plant Mol. Biol. Rep.* 5:387-405 (1987); Jefferson *et al.*, *EMBO J.* 6:3901-3907 (1987)); an R-locus gene, which encodes a product that regulates the production of anthocyanin pigments (red color) in plant tissues (Dellaporta *et al.*, *Stadler Symposium* 11:263-282 (1988)); a  $\beta$ -lactamase gene (Sutcliffe *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 75:3737-3741 (1978)), a gene which encodes an enzyme for which various chromogenic substrates are known (e.g., PADAC, a chromogenic cephalosporin); a luciferase gene (Ow *et al.*, *Science* 234:856-859 (1986)) a xyle gene (Zukowsky *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 80:1101-1105 (1983)) which encodes a catechol dioxygenase that can convert chromogenic catechols; an  $\alpha$ -amylase gene (Ikata *et al.*, *Bio/Technol.* 8:241-242 (1990)); a tyrosinase gene (Katz *et al.*, *J. Gen. Microbiol.* 129:2703-2714 (1983)) which encodes an enzyme capable of oxidizing tyrosine to DOPA and dopaquinone which in turn condenses to melanin; an  $\alpha$ -galactosidase, which will turn a chromogenic  $\alpha$ -galactose substrate.

Included within the terms "selectable or screenable marker genes" are also genes which encode a secretable marker whose secretion can be detected as a means of identifying or selecting for transformed cells. Examples include markers which encode a secretable antigen that can be identified by antibody interaction, or even secretable enzymes which can be detected catalytically. Secretable proteins fall into a number of classes, including small, diffusible proteins detectable, *e.g.*, by ELISA, small active enzymes detectable in extracellular solution (*e.g.*,  $\alpha$ -amylase,  $\beta$ -lactamase, phosphinothricin transferase), or proteins which are inserted or trapped in the cell wall (such as proteins which include a leader sequence such as that found in the expression unit of extension or tobacco PR-S). Other possible selectable and/or screenable marker genes will be apparent to those of skill in the art.

There are many methods for introducing nucleic acid molecules into plant cells. Suitable methods are believed to include virtually any method by which nucleic acid molecules may be introduced into a cell, such as by *Agrobacterium* infection or direct delivery of nucleic acid molecules such as, for example, by PEG-mediated transformation, by electroporation or by acceleration of DNA coated particles, etc. (Potrykus, *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 42:205-225 (1991); Vasil, *Plant Mol. Biol.* 25:925-937 (1994). For example, electroporation has been used to transform maize protoplasts (Fromm *et al.*, *Nature* 312:791-793 (1986)).

Technology for introduction of DNA into cells is well known to those of skill in the art. Four general methods for delivering a gene into cells have been described: (1) chemical methods (Graham and van der Eb, *Virology* 54:536-539 (1973)); (2) physical methods such as microinjection (Capecchi, *Cell* 22:479-488 (1980), electroporation (Wong and Neumann, *Biochem. Biophys. Res. Commun.* 107:584-587 (1982); Fromm *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 82:5824-5828 (1985); U. S. Patent No. 5,384,253; and the gene gun (Johnston and Tang, *Methods Cell Biol.* 43:353-365 (1994); (3) viral vectors (Clapp, *Clin. Perinatol.* 20:155-168 (1993); Lu *et al.*, *J. Exp. Med.* 178:2089-2096 (1993); Eglitis and Anderson, *Biotechnology* 6:608-614 (1988)); and (4) receptor-mediated mechanisms (Curiel *et al.*, *Hum. Gen. Ther.* 3:147-154 (1992); Wagner *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 89:6099-6103 (1992)).

Acceleration methods that may be used include, for example, microprojectile bombardment and the like. One example of a method for delivering transforming nucleic acid molecules to plant cells is microprojectile bombardment. This method has been reviewed by Yang and Christou, eds., *Particle Bombardment Technology for Gene Transfer*, Oxford Press, Oxford, England (1994)). Non-biological particles (microprojectiles) that may be coated with nucleic acids and delivered into cells by a propelling force. Exemplary particles include those comprised of tungsten, gold, platinum, and the like.

A particular advantage of microprojectile bombardment, in addition to it being an effective means of reproducibly, and stably transforming monocotyledons, is that neither the isolation of protoplasts

(Cristou *et al.*, *Plant Physiol.* 87:671-674 (1988)) nor the susceptibility of *Agrobacterium* infection is required. An illustrative embodiment of a method for delivering DNA into maize cells by acceleration is a biolistics alpha-particle delivery system, which can be used to propel particles coated with DNA through a screen, such as a stainless steel or Nytex screen, onto a filter surface covered with corn cells cultured in suspension. Gordon-Kamm *et al.*, describes the basic procedure for coating tungsten particles with DNA (*Plant Cell* 2:603-618 (1990)). The screen disperses the tungsten nucleic acid particles so that they are not delivered to the recipient cells in large aggregates. A particle delivery system suitable for use with the present invention is the helium acceleration PDS-1000/He gun which is available from Bio-Rad Laboratories (Bio-Rad, Hercules, CA) Sanford *et al.*, *Technique* 3:3-16 (1991)).

For the bombardment, cells in suspension may be concentrated on filters. Filters containing the cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the gun and the cells to be bombarded.

Alternatively, immature embryos or other target cells may be arranged on solid culture medium. The cells to be bombarded are positioned at an appropriate distance below the macroprojectile stopping plate. If desired, one or more screens are also positioned between the acceleration device and the cells to be bombarded. Through the use of techniques set forth herein one may obtain up to 1000 or more foci of cells transiently expressing a marker gene. The number of cells in a focus which express the exogenous gene product 48 hours post-bombardment often range from one to ten and average one to three.

In another alternative embodiment, plastids can be stably transformed. Methods suitable for plastid transformation in higher plants include particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination (Svab *et al.* *Proc. Natl. Acad. Sci. (U.S.A.)* 87:8526-8530 (1990); Svab and Maliga *Proc. Natl. Acad. Sci. (U.S.A.)* 90:913-917 (1993); Staub and Maliga, *EMBO J.* 12:601-606 (1993), U.S. Patents 5,451,513 and 5,545,818).

In bombardment transformation, one may optimize the prebombardment culturing conditions and the bombardment parameters to yield the maximum numbers of stable transformants. Both the physical and biological parameters for bombardment are important in this technology. Physical factors are those that involve manipulating the DNA/microprojectile precipitate or those that affect the flight and velocity of either the macro- or microprojectiles. Biological factors include all steps involved in manipulation of cells before and immediately after bombardment, the osmotic adjustment of target cells to help alleviate the trauma associated with bombardment, and also the nature of the transforming DNA, such as linearized DNA or intact supercoiled plasmids. It is believed that pre-bombardment manipulations are especially important for successful transformation of immature embryos.

Accordingly, it is contemplated that one may wish to adjust various aspects of the bombardment parameters in small scale studies to fully optimize the conditions. One may particularly wish to adjust physical parameters such as gap distance, flight distance, tissue distance, and helium pressure. One may also minimize the trauma reduction factors by modifying conditions which influence the physiological state of the recipient cells and which may therefore influence transformation and integration efficiencies. For example, the osmotic state, tissue hydration and the subculture stage or cell cycle of the recipient cells may be adjusted for optimum transformation. The execution of other routine adjustments will be known to those of skill in the art in light of the present disclosure.

*Agrobacterium*-mediated transfer is a widely applicable system for introducing genes into plant cells because the DNA can be introduced into whole plant tissues, thereby bypassing the need for regeneration of an intact plant from a protoplast. The use of *Agrobacterium*-mediated plant integrating vectors to introduce DNA into plant cells is well known in the art. See, for example, the methods described (Fraley *et al.*, *Biotechnology* 3:629-635 (1985); Rogers *et al.*, *Meth. Enzymol.* 153:253-277 (1987). Further, the integration of the Ti-DNA is a relatively precise process resulting in few rearrangements. The region of DNA to be transferred is defined by the border sequences, and intervening DNA is usually inserted into the plant genome as described (Spielmann *et al.*, *Mol. Gen. Genet.* 205:34 (1986)).

Modern *Agrobacterium* transformation vectors are capable of replication in *E. coli* as well as *Agrobacterium*, allowing for convenient manipulations as described (Klee *et al.*, *In: Plant DNA Infectious Agents*, Hohn and Schell, eds., Springer-Verlag, New York, pp. 179-203 (1985)). Moreover, recent technological advances in vectors for *Agrobacterium*-mediated gene transfer have improved the arrangement of genes and restriction sites in the vectors to facilitate construction of vectors capable of expressing various polypeptide coding genes. The vectors described have convenient multi-linker regions flanked by a promoter and a polyadenylation site for direct expression of inserted polypeptide coding genes and are suitable for present purposes (Rogers *et al.*, *Methods Enzymol.* 153:253-277 (1987)). In addition, *Agrobacterium* containing both armed and disarmed Ti genes can be used for the transformations. In those plant strains where *Agrobacterium*-mediated transformation is efficient, it is the method of choice because of the facile and defined nature of the gene transfer.

A transgenic plant formed using *Agrobacterium* transformation methods typically contains a single insert on one chromosome. Such transgenic plants can be referred to as being heterozygous for the added gene. More preferred is a transgenic plant that is homozygous for the added structural gene; *i.e.*, a transgenic plant that contains two added genes, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually mating (selfing) an

independent segregant transgenic plant that contains a single added gene, germinating some of the seed produced and analyzing the resulting plants produced for the gene of interest.

It is also to be understood that two different transgenic plants can also be mated to produce offspring that contain two independently segregating added, exogenous genes.

5 The present invention also provides for parts of the plants of the present invention. Plant parts, without limitation, include seed, endosperm, ovule and pollen. In a particularly preferred embodiment of the present invention, the plant part is a seed.

Transformation of plant protoplasts can be achieved using methods based on calcium phosphate precipitation, polyethylene glycol treatment, electroporation, and combinations of these treatments. See 10 for example (Potrykus *et al.*, *Mol. Gen. Genet* 205:193-200 (1986); Lorz *et al.*, *Mol. Gen. Genet.* 199:178, (1985); Fromm *et al.*, *Nature* 319:791,(1986); Uchimiya *et al.*, *Mol. Gen. Genet.* 204:204 (1986); Callis *et al.*, *Genes and Development* 1183 (1987); Marcotte *et al.*, *Nature* 335:454 (1988)).

Application of these systems to different plant strains depends upon the ability to regenerate that particular plant strain from protoplasts. Illustrative methods for the regeneration of cereals from 15 protoplasts are described (Fujimura *et al.*, *Plant Tissue Culture Letters* 2:74 (1985); Toriyama *et al.*, *Theor Appl. Genet.* 205:34 (1986); Yamada *et al.*, *Plant Cell Rep.* 4:85 (1986); Abdullah *et al.*, *Biotechnology* 4:1087 (1986)).

To transform plant strains that cannot be successfully regenerated from protoplasts, other ways to introduce DNA into intact cells or tissues can be utilized. For example, regeneration of cereals from 20 immature embryos or explants can be effected as described (Vasil, *Biotechnology* 6: 397 (1988)). In addition, "particle gun" or high-velocity microprojectile technology can be utilized (Vasil *et al.*, *Bio/Technology* 10:667, (1992)).

Using the latter technology, DNA is carried through the cell wall and into the cytoplasm on the surface of small metal particles as described (Klein *et al.*, *Nature* 328:70 (1987); Klein *et al.*, *Proc. Natl. 25 Acad. Sci. (U.S.A.)* 85:8502-8505 (1988); McCabe *et al.*, *Biotechnology* 6 :923 (1988)). The metal particles penetrate through several layers of cells and thus allow the transformation of cells within tissue explants.

Other methods of cell transformation can also be used and include but are not limited to introduction of DNA into plants by direct DNA transfer into pollen (Hess *et al.*, *Intern Rev. Cytol.* 30 107:367 (1987); Luo *et al.*, *Plant Mol. Biol.. Reporter* 6:165 (1988)), by direct injection of DNA into reproductive organs of a plant (Pena *et al.*, *Nature* 325:274 (1987)), or by direct injection of DNA into the cells of immature embryos followed by the rehydration of desiccated embryos (Neuhaus *et al.*, *Theor. Appl. Genet.* 75:30 (1987)).

The regeneration, development, and cultivation of plants from single plant protoplast transformants or from various transformed explants is well known in the art (Weissbach and Weissbach, In: *Methods for Plant Molecular Biology*, (Eds.), Academic Press, Inc. San Diego, CA, (1988)). This regeneration and growth process typically includes the steps of selection of transformed cells, culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. Transgenic embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil.

The development or regeneration of plants containing the foreign, exogenous gene that encodes a protein of interest is well known in the art. Preferably, the regenerated plants are self-pollinated to provide homozygous transgenic plants, as discussed before. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the present invention containing a desired polypeptide is cultivated using methods well known to one skilled in the art.

There are a variety of methods for the regeneration of plants from plant tissue. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated.

Methods for transforming dicots, primarily by use of *Agrobacterium tumefaciens*, and obtaining transgenic plants have been published for cotton (U. S. Patent No. 5,004,863; U.S. Patent No. 5,159,135; U.S. Patent No. 5,518,908); soybean (U. S. Patent No. 5,569,834; U. S. Patent No. 5,416,011; McCabe *et al.*, *Biotechnology* 6:923 (1988), Christou *et al.*, *Plant Physiol.* 87:671-674 (1988)); *Brassica* (U. S. Patent No. 5,463,174); peanut (Cheng *et al.*, *Plant Cell Rep.* 15:653-657 (1996), McKently *et al.*, *Plant Cell Rep.* 14:699-703 (1995)); papaya (Yang *et al.*, (1996)); pea (Grant *et al.*, *Plant Cell Rep.* 15:254-258, (1995)).

Transformation of monocotyledons using electroporation, particle bombardment, and *Agrobacterium* have also been reported. Transformation and plant regeneration have been achieved in asparagus (Bytebier *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 84:5345 (1987)); barley (Wan and Lemaux, *Plant Physiol* 104:37 (1994)); maize (Rhodes *et al.*, *Science* 240:204 (1988), Gordon-Kamm *et al.*, *Plant Cell* 2:603 (1990), Fromm *et al.*, *Bio/Technology* 8:833 (1990), Koziel *et al.*, *Bio/Technology* 11:194 (1993), Armstrong *et al.*, *Crop Science* 35:550-557 (1995)); oat (Somers *et al.*, *Bio/Technology* 10:1589 (1992)); orchardgrass (Horn *et al.*, *Plant Cell Rep.* 7:469 (1988)); rice (Toriyama *et al.*, *Theor Appl. Genet.* 205:34 (1986); Park *et al.*, *Plant Mol. Biol.* 32:1135-1148 (1996); Abedinia *et al.*, *Aust. J. Plant Physiol.* 24:133-141 (1997); Zhang and Wu, *Theor. Appl. Genet.* 76:835 (1988); Zhang *et al.* *Plant Cell*

Rep. 7:379 (1988); Battraw and Hall, *Plant Sci.* 86:191-202 (1992); Christou *et al.*, *Bio/Technology* 9:957 (1991)); sugarcane (Bower and Birch, *Plant J.* 2:409 (1992)); tall fescue (Wang *et al.*, *Bio/Technology* 10:691 (1992)), and wheat (Vasil *et al.*, *Bio/Technology* 10:667 (1992); U. S. Patent No. 5,631,152.

Assays for gene expression based on the transient expression of cloned nucleic acid constructs have been developed by introducing the nucleic acid molecules into plant cells by polyethylene glycol treatment, electroporation, or particle bombardment (Marcotte *et al.*, *Nature* 335:454-457 (1988); Marcotte *et al.*, *Plant Cell* 1:523-532 (1989); McCarty *et al.*, *Cell* 66:895-905 (1991); Hattori *et al.*, *Genes Dev.* 6:609-618 (1992); Goff *et al.*, *EMBO J.* 9:2517-2522 (1990)). Transient expression systems may be used to functionally dissect gene constructs (See generally, Mailga *et al.*, *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1995)).

Any of the nucleic acid molecules of the present invention may be introduced into a plant cell in a permanent or transient manner in combination with other genetic elements such as vectors, promoters, enhancers etc. Further any of the nucleic acid molecules encoding a *B. thuringiensis* protein or fragment thereof of the present invention may be introduced into a plant cell in a manner that allows for over expression of the protein or fragment thereof encoded by the nucleic acid molecule.

Antibodies have been expressed in plants (Hiatt *et al.*, *Nature* 342:76-78 (1989); Conrad and Fielder, *Plant Mol. Biol.* 26:1023-1030 (1994)). Cytoplasmic expression of a scFv (single-chain Fv antibodies) has been reported to delay infection by artichoke mottled crinkle virus. Transgenic plants that express antibodies directed against endogenous proteins may exhibit a physiological effect (Philips *et al.*, *EMBO J.* 16:4489-4496 (1997); Marion-Poll, *Trends in Plant Science* 2:447-448 (1997)). For example, expressed anti-abscisic antibodies reportedly result in a general perturbation of seed development (Philips *et al.*, *EMBO J.* 16:4489-4496 (1997)).

Antibodies that are catalytic may also be expressed in plants (abzymes). The principle behind abzymes is that since antibodies may be raised against many molecules, this recognition ability can be directed toward generating antibodies that bind transition states to force a chemical reaction forward (Persidas, *Nature Biotechnology* 15:1313-1315 (1997); Baca *et al.*, *Ann. Rev. Biophys. Biomol. Struct.* 26:461-493 (1997)). The catalytic abilities of abzymes may be enhanced by site directed mutagenesis. Examples of abzymes are, for example, set forth in U.S. Patent No. 5,658,753; U.S. Patent No. 5,632,990; U.S. Patent No. 5,631,137; U.S. Patent 5,602,015; U.S. Patent No. 5,559,538; U.S. Patent No. 5,576,174; U.S. Patent No. 5,500,358; U.S. Patent 5,318,897; U.S. Patent No. 5,298,409; U.S. Patent No. 5,258,289 and U.S. Patent No. 5,194,585, all of which are herein incorporated in their entirety.

It is understood that any of the antibodies of the present invention may be expressed in plants and that such expression can result in a physiological effect. It is also understood that any of the expressed antibodies may be catalytic.

## 5 Exemplary Uses of the Agents of the Present Invention

Nucleic acid molecules and fragments thereof of the present invention may be employed to obtain other nucleic acid molecules from the same species. Such nucleic acid molecules include the nucleic acid molecules that encode the complete coding sequence of a protein and promoters and flanking sequences of such molecules. In addition, such nucleic acid molecules include nucleic acid molecules  
10 that encode for other isozymes or gene family members. Such molecules can be readily obtained by using the above-described nucleic acid molecules or fragments thereof to screen genomic libraries obtained from *B. thuringiensis*. Methods for forming such libraries are well known in the art.

Nucleic acid molecules and fragments thereof of the present invention may also be employed to obtain other nucleic acid molecules such as nucleic acid homologues. Such homologues include the  
15 nucleic acid nucleic acid homologues of non-*B. thuringiensis* species including the nucleic acid molecules that encode, in whole or in part, protein homologues of other species or other organisms, sequences of genetic elements such as promoters and transcriptional regulatory elements. Such molecules can be readily obtained by using the above-described nucleic acid molecules or fragments thereof to screen cDNA or genomic libraries. Methods for forming such libraries are well known in the  
20 art. Such homologue molecules may differ in their nucleotide sequences from those found in one or more of SEQ ID NO: 1 through SEQ ID NO: 8283 or complements thereof because complete complementarity is not needed for stable hybridization. The nucleic acid molecules of the present invention therefore also include molecules that, although capable of specifically hybridizing with the nucleic acid molecules may lack "complete complementarity." In a particular embodiment, methods or 3' or 5' RACE may be used  
25 (Frohman, M.A. *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 85:8998-9002 (1988); Ohara, O. *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:5673-5677 (1989)) to obtain such sequences.

Any of a variety of methods may be used to obtain one or more of the above-described nucleic acid molecules (Zamechik *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 83:4143-4146 (1986); Goodchild *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 85: 5507-5511 (1988); Wickstrom *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*  
30 85:1028-1032 (1988); Holt *et al.*, *Molec. Cell. Biol.* 8:963-973 (1988); Gerwitz *et al.*, *Science* 242: 1303-1306 (1988); Anfossi *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:3379-3383 (1989); Becker *et al.*, *EMBO J.* 8:3685-3691 (1989)). Automated nucleic acid synthesizers may be employed for this purpose. In lieu of such synthesis, the disclosed nucleic acid molecules may be used to define a pair of primers that can be



used with the polymerase chain reaction (Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 51:263-273 (1986); Erlich *et al.*, European Patent 50,424; European Patent 84,796, European Patent 258,017; European Patent 237,362; Mullis, European Patent 201,184; Mullis *et al.*, U.S. Patent 4,683,202; Erlich, U.S. Patent 4,582,788; and Saiki *et al.*, U.S. Patent 4,683,194) to amplify and obtain any desired nucleic acid molecule or fragment.

The nucleic acid molecules of the present invention may be used for physical mapping. Physical mapping, in conjunction with linkage analysis, can enable the isolation of genes. Physical mapping has been reported to identify the markers closest in terms of genetic recombination to a gene target for cloning. Once a DNA marker is linked to a gene of interest, the chromosome walking technique can be used to find the genes via overlapping clones. For chromosome walking, random molecular markers or established molecular linkage maps are used to conduct a search to localize the gene adjacent to one or more markers. A chromosome walk (Bukanov and Berg, *Mo. Microbiol.* 11:509-523 (1994); Birkenbihl and Vielmetter *Nucleic Acids Res.* 17:5057-5069 (1989); Wenzel and Herrmann, *Nucleic Acids Res.* 16:8323-8336 (1988) is then initiated from the closest linked marker. Starting from the selected clones, labeled probes specific for the ends of the insert DNA are synthesized and used as probes in hybridizations against a representative library. Clones hybridizing with one of the probes are picked and serve as templates for the synthesis of new probes; by subsequent analysis, contigs are produced.

The degree of overlap of the hybridizing clones used to produce a contig can be determined by comparative restriction analysis. Comparative restriction analysis can be carried out in different ways all of which exploit the same principle; two clones of a library are very likely to overlap if they contain a limited number of restriction sites for one or more restriction endonucleases located at the same distance from each other. The most frequently used procedures are, fingerprinting (Coulson *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 83:7821-7821, (1986); Knott *et al.*, *Nucleic Acids Res.* 16:2601-2612 (1988); Eiglmeier *et al.*, *Mol. Microbiol.* 7:197-206 (1993), 1993), restriction fragment mapping (Smith and Birnstiel, *Nucleic Acids Res.* 3:2387-2398 (1976)); or the "landmarking" technique (Charlebois *et al.*, *J. Mol. Biol.* 222:509-524 (1991)).

It is understood that the nucleic acid molecules of the present invention may in one embodiment be used in physical mapping. In a preferred embodiment, nucleic acid molecules of the present invention may in one embodiment be used in the physical mapping of *B. thuringiensis*.

Nucleic acid molecules of the present invention can be used in comparative mapping. Comparative mapping within families provides a method to assess the degree of sequence conservation, gene order, ploidy of species, ancestral relationships and the rates at which individual genomes are evolving. Comparative mapping has been carried out by cross-hybridizing molecular markers across

species within a given family. As in genetic mapping, molecular markers are needed but instead of direct hybridization to mapping filters, the markers are used to select large insert clones from a total genomic DNA library of a related species. The selected clones, each a representative of a single marker, can then be used to physically map the region in the target species. The advantage of this method for comparative mapping is that no mapping population or linkage map of the target species is needed and the clones may also be used in other closely related species. By comparing the results obtained by genetic mapping in model organisms, with those from other species, similarities of genomic structure among species can be established. Cross-hybridization of RFLP markers has been reported and conserved gene order has been established in many studies. Such macroscopic synteny is utilized for the estimation of correspondence of loci among these organisms. It is understood that markers of the present invention may in another embodiment be used in comparative mapping. In a preferred embodiment the markers of present invention may be used in the comparative mapping of spore-forming Gram-positive bacteria.

In an aspect of the present invention, one or more of the agents of the present invention may be used to detecting the presence, absence or level of an organism, preferably a *Bacillus*, and more preferably a *B. thuringiensis* in a sample. In another aspect of the present invention, one or more of the nucleic molecules of the present invention are used to determine the level (i.e., the concentration of mRNA in a sample, etc.) or pattern (i.e., the kinetics of expression, rate of decomposition, stability profile, etc.) of the expression encoded in part or whole by one or more of the nucleic acid molecule of the present invention (collectively, the "Expression Response" of a cell or tissue). As used herein, the Expression Response manifested by a cell or tissue is said to be "altered" if it differs from the Expression Response of cells or tissues of organisms not exhibiting the phenotype. To determine whether a Expression Response is altered, the Expression Response manifested by the cell or tissue of the organism exhibiting the phenotype is compared with that of a similar cell or tissue sample of a organism not exhibiting the phenotype. As will be appreciated, it is not necessary to re-determine the Expression Response of the cell or tissue sample of organisms not exhibiting the phenotype each time such a comparison is made; rather, the Expression Response of a particular organism may be compared with previously obtained values of normal organism. As used herein, the phenotype of the organism is any of one or more characteristics of an organism.

Nucleic acid molecules of the present invention can be used to monitor expression. A microarray-based method for high-throughput monitoring of gene expression may be utilized to measure gene-specific hybridization targets. This 'chip'-based approach involves using microarrays of nucleic acid molecules as gene-specific hybridization targets to quantitatively measure expression of the corresponding genes (Schena *et al.*, *Science* 270:467-470 (1995); Shalon, Ph.D. Thesis, Stanford

University (1996)). Every nucleotide in a large sequence can be queried at the same time. Hybridization can be used to efficiently analyze nucleotide sequences.

Several microarray methods have been described. One method compares the sequences to be analyzed by hybridization to a set of oligonucleotides or cDNA molecules representing all possible subsequences (Bains and Smith, *J. Theor. Biol.* 135:303 (1989)). A second method hybridizes the sample to an array of oligonucleotide or cDNA probes. An array consisting of oligonucleotides or cDNA molecules complementary to subsequences of a target sequence can be used to determine the identity of a target sequence, measure its amount, and detect differences between the target and a reference sequence. Nucleic acid molecules microarrays may also be screened with protein molecules or fragments thereof to determine nucleic acid molecules that specifically bind protein molecules or fragments thereof.

The microarray approach may also be used with polypeptide targets (U.S. Patent No. 5,445,934; U.S. Patent No. 5,143,854; U.S. Patent No. 5,079,600; U.S. Patent No. 4,923,901). Essentially, polypeptides are synthesized on a substrate (microarray) and these polypeptides can be screened with either protein molecules or fragments thereof or nucleic acid molecules in order to screen for either protein molecules or fragments thereof or nucleic acid molecules that specifically bind the target polypeptides (Fodor *et al.*, *Science* 251:767-773 (1991)).

It is understood that one or more of the molecules of the present invention, preferably one or more of the nucleic acid molecules or protein molecules or fragments thereof of the present invention may be utilized in a microarray based method. In a preferred embodiment of the present invention, one or more of the *B. thuringiensis* nucleic acid molecules or protein molecules or fragments thereof of the present invention may be utilized in a microarray based method. A particular preferred microarray embodiment of the present invention is a microarray comprising nucleic acid molecules encoding genes or fragments thereof that are homologues of known genes or nucleic acid molecules that comprise genes or fragments thereof that elicit only limited or no matches to known genes. A further preferred microarray embodiment of the present invention is a microarray comprising nucleic acid molecules having genes or fragments thereof that are homologues of known genes and nucleic acid molecules that comprise genes or fragment thereof that elicit only limited or no matches to known genes.

In a preferred embodiment, the microarray of the present invention comprises at least 10 nucleic acid molecules that specifically hybridize under high stringency to at least 10 nucleic acid molecules encoding *B. thuringiensis* protein or fragments thereof set forth in Table 1. In a more preferred embodiment, the microarray of the present invention comprises at least 100 nucleic acid molecules that specifically hybridize under high stringency to at least 100 nucleic acid molecules that encode a *B. thuringiensis* protein or fragment thereof set forth in Table 1. In an even more preferred embodiment, the

microarray of the present invention comprises at least 1,000 nucleic acid molecules that specifically hybridize under high stringency to at least 1,000 nucleic acid molecules that encode a *B. thuringiensis* protein or fragment thereof set forth in Table 1. In a further even more preferred embodiment, the microarray of the present invention comprises at least 2,500 nucleic acid molecules that specifically hybridize under high stringency to at least 2,500 nucleic acid molecules that encode a *B. thuringiensis* protein or fragment thereof set forth in Table 1. While it is understood that a single nucleic acid molecule may encode more than one protein homologue or fragment thereof, in a preferred embodiment, at least 50%, preferably at least 70%, more preferably at least 80%, even more preferably at least 90% of the nucleic acid molecules that comprise the microarray contain one protein or fragment thereof.

In a preferred embodiment, the microarray of the present invention comprises at least 10 nucleic acid molecules that specifically hybridize under high stringency to at least 10 nucleic acid molecules selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283 or fragment thereof or complement of either. In a more preferred embodiment, the microarray of the present invention comprises at least 100 nucleic acid molecules that specifically hybridize under high stringency to at least 100 nucleic acid molecules that encode a *B. thuringiensis* protein or fragment thereof set forth in Table 1. In an even more preferred embodiment, the microarray of the present invention comprises at least 1,000 nucleic acid molecules that specifically hybridize under high stringency to at least 1,000 nucleic acid molecules that encode a *B. thuringiensis* protein or fragment thereof set forth in Table 1. In a further even more preferred embodiment, the microarray of the present invention comprises at least 2,500 nucleic acid molecules that specifically hybridize under high stringency to at least 2,500 nucleic acid molecules that encode a *B. thuringiensis* protein or fragment thereof set forth in Table 1. While it is understood that a single nucleic acid molecule may encode more than one protein homologue or fragment thereof, in a preferred embodiment, at least 50%, preferably at least 70%, more preferably at least 80%, even more preferably at least 90% of the nucleic acid molecules that comprise the microarray contain one protein homologue or fragment thereof.

Nucleic acid molecules of the present invention may be used in site directed mutagenesis. Site-directed mutagenesis may be utilized to modify nucleic acid sequences, particularly as it is a technique that allows one or more of the amino acids encoded by a nucleic acid molecule to be altered (e.g. a threonine to be replaced by a methionine). Three basic methods for site-directed mutagenesis are often employed. These are cassette mutagenesis (Wells *et al.*, *Gene* 34:315-23 (1985)); primer extension (Gilliam *et al.*, *Gene* 12:129-137 (1980)); Zoller and Smith, *Methods Enzymol.* 100:468-500 (1983); and Dalbadie-McFarland *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 79:6409-6413 (1982)) and methods based upon PCR (Scharf *et al.*, *Science* 233:1076-1078 (1986); Higuchi *et al.*, *Nucleic Acids Res.* 16:7351-7367

(1988)). Site-directed mutagenesis approaches are also described in European Patent 0 385 962, European Patent 0 359 472, and PCT Patent Application WO 93/07278.

Site-directed mutagenesis strategies have been applied to plants for both *in vitro* as well as *in vivo* site-directed mutagenesis (Lanz *et al.*, *J. Biol. Chem.* 266:9971-9976 (1991); Kovgan and Zhdanov, *Biokhnologiya* 5: 148-154, No. 207160n, Chemical Abstracts 110: 225 (1989); Ge *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:4037-4041 (1989), Zhu *et al.*, *J. Biol. Chem.* 271:18494-18498 (1996), Chu *et al.*, *Biochemistry* 33:6150-6157 (1994), Small *et al.*, *EMBO J.* 11:1291-1296 (1992), Cho *et al.*, *Mol. Biotechnol.* 8:13-16 (1997), Kita *et al.*, *J. Biol. Chem.* 271:26529-26535 (1996), Jin *et al.*, *Mol. Microbiol.* 7:555-562 (1993), Hatfield and Vierstra, *J. Biol. Chem.* 267:14799-14803 (1992), Zhao *et al.*, *Biochemistry* 31:5093-5099 (1992)).

Any of the nucleic acid molecules of the present invention may either be modified by site-directed mutagenesis or used as, for example, nucleic acid molecules that are used to target other nucleic acid molecules for modification. It is understood that mutants with more than one altered nucleotide can be constructed using techniques that practitioners skilled in the art are familiar with such as isolating restriction fragments and ligating such fragments into an expression vector (*see*, for example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989)). In a preferred embodiment of the present invention, one or more of the nucleic acid molecules or fragments thereof of the present invention may be modified by site-directed mutagenesis.

In addition to the above discussed procedures, practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of macromolecules (e.g., DNA molecules, plasmids, etc.), generation of recombinant organisms and the screening and isolating of clones, (*see* for example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989); Mailga *et al.*, *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1995); Birren *et al.*, *Genome Analysis: Analyzing DNA*, 1, Cold Spring Harbor, New York).

#### (a) Computer Media

The nucleotide sequence provided in SEQ ID NO: 1, through SEQ ID NO: 8283 or fragment thereof, or complement thereof, or a nucleotide sequence at least 90% identical, preferably 95%, identical even more preferably 99% or 100% identical to the sequence provided in SEQ ID NO: 1 through SEQ ID NO: 8283 or fragment thereof, or complement thereof, can be "provided" in a variety of mediums to facilitate use. Such a medium can also provide a subset thereof in a form that allows a skilled artisan to examine the sequences.

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium that can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc, storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention.

As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate media comprising the nucleotide sequence information of the present invention. A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing one or more of nucleotide sequences of the present invention, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990)) and BLAZE (Brutlag *et al.*, *Comp. Chem.* 17:203-207 (1993)) search algorithms on a Sybase system can be used to identify open reading frames (ORFs) within the genome that contain homology to ORFs or proteins from other organisms. Such ORFs are protein-encoding fragments within the sequences of the present invention and are useful in producing commercially important proteins such as enzymes used in amino acid biosynthesis, metabolism, transcription, translation, RNA processing, nucleic acid and a protein degradation, protein modification, and DNA replication, restriction, modification, recombination, and repair.

The present invention further provides systems, particularly computer-based systems, which contain the sequence information described herein. Such systems are designed to identify commercially important fragments of the nucleic acid molecule of the present invention. As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based system are suitable for use in the present invention.

As indicated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory that can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention. As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequence of the present invention that match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are available can be used in the computer-based systems of the present invention. Examples of such software include, but are not limited to, MacPattern (EMBL), BLASTIN and BLASTIX (NCBIA). One of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems.

The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that during searches for commercially important fragments of the nucleic acid molecules of the present invention, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequences the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzymatic active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, cis elements, hairpin structures and inducible expression elements (protein binding sequences).

Thus, the present invention further provides an input means for receiving a target sequence, a data storage means for storing the target sequences of the present invention sequence identified using a search means as described above, and an output means for outputting the identified homologous sequences. A variety of structural formats for the input and output means can be used to input and output information in the computer-based systems of the present invention. A preferred format for an output means ranks fragments of the sequence of the present invention by varying degrees of homology to the target sequence or target motif. Such presentation provides a skilled artisan with a ranking of sequences which contain various amounts of the target sequence or target motif and identifies the degree of homology contained in the identified fragment.

A variety of comparing means can be used to compare a target sequence or target motif with the data storage means to identify sequence fragments sequence of the present invention. For example, implementing software which implement the BLAST and BLAZE algorithms (Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990)) can be used to identify open frames within the nucleic acid molecules of the present invention. A skilled artisan can readily recognize that any one of the publicly available homology search programs can be used as the search means for the computer-based systems of the present invention.

The present invention also provides, in another aspect, a method for identifying one or more genes encoding insect inhibitory proteins in the sequences of one or more plasmids of a *Bacillus thuringiensis*, wherein the method of the present invention comprises the steps of: a) isolating and purifying plasmid DNA; b) constructing a DNA library from the isolated and purified plasmid DNA; c) sequencing the DNA library to obtain a set of plasmid DNA sequences; d) comparing the set of DNA sequences with a set of chromosomal DNA sequences, wherein the set of chromosomal DNA sequences comprises sequences depicted in SEQ ID No: 1 through SEQ ID No: 8283; e) identify common sequences, which are identified both in the set of plasmid DNA sequences and in the set of chromosomal DNA sequences; f) subtracting the common sequences from the set of plasmid DNA sequences to obtain a subtracted set of plasmid DNA sequences; g) assembling the subtracted set of DNA sequences to obtain contigs and sequences; h) determining open reading frames in the contigs and sequences; and h) identifying one or more genes encoding insect inhibitory proteins in the sequences of one or more plasmids of said *Bacillus thuringiensis*.

Methods for isolating and purifying plasmid DNA are well known in the art. For example, DNA samples comprising endogenous plasmids from *Bacillus thuringiensis* strains can be isolated and purified using the method of Hansen and Olsen (*J. Bacteriology*, Vol. 135, pp. 227-238 (1978)). Methods for constructing a DNA library are well known in the art.



As used herein, "a contig" refers to a contiguous nucleotide sequence that is assembled from two or more constituent nucleotide sequences that share common or overlapping regions of sequence homology. For example, the nucleotide sequences of two or more nucleic acid fragments can be compared and aligned in order to identify common or overlapping sequences. Where common or overlapping sequences exist between two or more nucleic acid fragments, the sequences (and thus their corresponding nucleic acid fragments) can be assembled into a single contiguous nucleotide sequence.

It is well known to a person skilled in the art that the sequence data from a large scale shotgun sequencing project can be processed and assembled into contigs, which represent a reconstruction of the original chromosomal genome sequence from the cloned fragments. Programs are available in the public domain that can analyze the sequence output and assemble the sequences into larger sequence regions representing contiguous sequences of the target genome. Examples of such programs can be found at, for example, <http://genome.wustl.edu/gsc>, <http://www.sanger.ac.uk>, and <http://www.mbt.washington.edu>. An example of a sequence reading program is Phred (<http://www.mbt.washington.edu>). Phred reads DNA sequencer trace data, calls bases, assigns quality values to the bases, and writes the base calls and quality values to output files.

The process of assembling DNA sequence fragments generally involves three phases; the overlap phase, the layout phase and the multi-alignment, or consensus, phase. In the overlap phase, each fragment is compared against every other fragment to determine if they share a common subsequence, an indication that they were potentially sampled from overlapping stretches of the original DNA strand. Pairs of fragments are compared in two ways; 1) with both fragments in the same relative orientation, and 2) with one of the fragments having been reverse complemented. In the layout phase, a series of alternate assemblies or layouts of the fragments based on the pairwise overlaps is generated. A layout specifies the relative locations and orientations of the fragments with respect to each other and is typically visualized as an arrangement of overlapping directed lines, one for each fragment. The general criterion for the layout phase is to produce plausible assemblies of maximum likelihood. In this manner, it can be determined if there is more than one way to put the pieces together and if different solutions appear equally plausible. In such a case, one would return to the lab and obtain additional information to resolve the ambiguity. The multi-alignment, or consensus, phase uses more information than just the pairwise alignments in the layout. The sequences of all the fragments in a layout are simultaneously aligned, giving a final set of contigs representing regions of the target genome. An example of an assembly program is PHRAP, which can be found at <http://chimera.biotech.washington.edu/UWGC/tools/phrap.htm>.

A characteristic feature of a DNA sequence is that it can be compared with other known DNA sequences. Sequence comparisons can be undertaken by determining the similarity of the test or query sequence with sequences in publicly available or propriety databases ("similarity analysis") or by searching for certain motifs ("intrinsic sequence analysis") (e.g. *cis* elements) (Coulson, *Trends in Biotechnology* 12:76-80 (1994); Birren, *et al.*, *Genome Analysis* 1:543-559 (1997)).

Similarity analysis includes database search and alignment. Examples of public databases include the DNA Database of Japan (DDBJ) (<http://www.ddbj.nig.ac.jp/>); Genebank (<http://www.ncbi.nlm.nih.gov/web/Genbank/Index.html>); and the European Molecular Biology Laboratory Nucleic Acid Sequence Database (EMBL) ([http://www.ebi.ac.uk/ebi\\_docs/embl\\_db.html](http://www.ebi.ac.uk/ebi_docs/embl_db.html)).

A number of different search algorithms have been developed, one example of which are the suite of programs referred to as BLAST programs. There are five implementations of BLAST, three designed for nucleotide sequences queries (BLASTN, BLASTX, and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, *Trends in Biotechnology* 12:76-80 (1994); Birren *et al.*, *Genome Analysis* 1:543-559 (1997)).

BLASTN takes a nucleotide sequence (the query sequence) and its reverse complement and searches them against a nucleotide sequence database. BLASTN was designed for speed, not maximum sensitivity, and may not find distantly related coding sequences. BLASTX takes a nucleotide sequence, translates it in three forward reading frames and three reverse complement reading frames, and then compares the six translations against a protein sequence database. BLASTX is useful for sensitive analysis of preliminary (single-pass) sequence data and is tolerant of sequencing errors (Gish and States, *Nature Genetics* 3:266-272 (1993)). BLASTN and BLASTX may be used in concert for analyzing sequence data (Coulson, *Trends in Biotechnology* 12:76-80 (1994); Birren *et al.*, *Genome Analysis* 1:543-559 (1997)).

Given a nucleotide coding sequence and the predicted protein which may be produced from that sequence, it is often preferable to use the protein as the query sequence to search a database because of the greatly increased sensitivity to detect more subtle relationships. This is due to the larger alphabet of proteins (20 amino acids) compared with the alphabet of nucleic acid sequences (4 bases), where it is far easier to obtain a match by chance. In addition, with nucleotide alignments, only a match (positive score) or a mismatch (negative score) is obtained, but with proteins, the presence of conservative amino acid substitutions can be taken into account. Here, a mismatch may yield a positive score if the non-identical residue has physical/chemical properties similar to the one with which it aligns. Various scoring matrices are used to supply the substitution scores of all possible amino acid alignments. A general purpose scoring system is the BLOSUM62 matrix (Henikoff and Henikoff, *Proteins* 17:49-61 (1993)), which is

currently the default choice for BLAST programs. BLOSUM62 is tailored for alignments of moderately diverged sequences and thus may not yield the best results under all conditions. Altschul, *J. Mol. Biol.* 36:290-300 (1993), uses a combination of three matrices to cover all contingencies. This may improve sensitivity, but at the expense of slower searches. In practice, a single BLOSUM62 matrix is often used but others (PAM40 and PAM250) may be attempted when additional analysis is necessary. Low PAM matrices are directed at detecting very strong but localized sequence similarities, whereas high PAM matrices are directed at detecting long but weak alignments between very distantly related sequences.

Homologues in other organisms are available that can be used for comparative sequence analysis. Multiple alignments are performed to study similarities and differences in a group of related sequences.

CLUSTAL W is a multiple sequence alignment package that performs progressive multiple sequence alignments based on the method of Feng and Doolittle, *J. Mol. Evol.* 25:351-360 (1987). Each pair of sequences is aligned and the distance between each pair is calculated; from this distance matrix, a guide tree is calculated, and all of the sequences are progressively aligned based on this tree. A feature of the program is its sensitivity to the effect of gaps on the alignment; gap penalties are varied to encourage the insertion of gaps in probable loop regions instead of in the middle of structured regions. Users can specify gap penalties, choose between a number of scoring matrices, or supply their own scoring matrix for both the pairwise alignments and the multiple alignments. CLUSTAL W for UNIX and VMS systems is available at: <ftp://ebi.ac.uk>. Another program is MACAW (Schuler *et al.*, *Proteins, Struct. Func. Genet.* 9:180-190 (1991), for which both Macintosh and Microsoft Windows versions are available. MACAW uses a graphical interface, provides a choice of several alignment algorithms, and is available by anonymous ftp at: [ncbi.nlm.nih.gov \(directory/pub/macaw\)](ftp://ncbi.nlm.nih.gov/directory/pub/macaw).

Sequence motifs are derived from multiple alignments and can be used to examine individual sequences or an entire database for subtle patterns. With motifs, it is sometimes possible to detect distant relationships that may not be demonstrable based on comparisons of primary sequences alone. Currently, the largest collection of sequence motifs in the world is PROSITE (Bairoch and Bucher, *Nucleic Acid Research* 22:3583-3589 (1994)). PROSITE may be accessed via either the ExPASy server on the World Wide Web or anonymous ftp site. Many commercial sequence analysis packages also provide search programs that use PROSITE data.

A resource for searching protein motifs is the BLOCKS E-mail server developed by S. Henikoff, *Trends Biochem Sci.* 18:267-268 (1993); Henikoff and Henikoff, *Nucleic Acid Research* 19:6565-6572 (1991); Henikoff and Henikoff, *Proteins* 17:49-61 (1993). BLOCKS searches a protein or protein encoding nucleotide sequence against a database of protein motifs or "blocks." Blocks are defined as short, ungapped multiple alignments that represent highly conserved protein patterns. The blocks

themselves are derived from entries in PROSITE as well as other sources. Either a protein or nucleotide query can be submitted to the BLOCKS server; if a nucleotide sequence is submitted, the sequence is translated in all six reading frames and motifs are sought in these conceptual translations. Once the search is completed, the server will return a ranked list of significant matches, along with an alignment of the query sequence to the matched BLOCKS entries.

Conserved protein domains can be represented by two-dimensional matrices, which measure either the frequency or probability of the occurrences of each amino acid residue and deletions or insertions in each position of the domain. This type of model, when used to search against protein databases, is sensitive and usually yields more accurate results than simple motif searches. Two popular implementations of this approach are profile searches (such as GCG program ProfileSearch) and Hidden Markov Models (HMMs) (Krough *et al.*, *J. Mol. Biol.* 235:1501-1531 (1994); Eddy, *Current Opinion in Structural Biology* 6:361-365 (1996)). In both cases, a large number of common protein domains have been converted into profiles, as present in the PROSITE library, or HMM models, as in the Pfam protein domain library (Sonnhammer *et al.*, *Proteins* 28:405-420 (1997)). Pfam contains more than 500 HMM models for enzymes, transcription factors, signal transduction molecules, and structural proteins. Protein databases can be queried with these profiles or HMM models, which will identify proteins containing the domain of interest. For example, HMMSW or HMMFS, two programs in a public domain package called HMMER (Sonnhammer *et al.*, *Proteins* 28:405-420 (1997)) can be used.

PROSITE and BLOCKS represent collected families of protein motifs. Thus, searching these databases entails submitting a single sequence to determine whether or not that sequence is similar to the members of an established family. Programs working in the opposite direction compare a collection of sequences with individual entries in the protein databases. An example of such a program is the Motif Search Tool, or MoST (Tatusov *et al.*, *Proc. Natl. Acad. Sci.* 91:12091-12095 (1994).) On the basis of an aligned set of input sequences, a weight matrix is calculated by using one of four methods (selected by the user); a weight matrix is simply a representation, position by position in an alignment, of how likely a particular amino acid will appear. The calculated weight matrix is then used to search the databases. To increase sensitivity, newly found sequences are added to the original data set, the weight matrix is recalculated, and the search is performed again. This procedure continues until no new sequences are found.

The present invention also further provides, in another aspect, a method for identifying plasmid DNA sequences of a *Bacillus* species, the method comprising the steps of a) identifying a *Bacillus* species strain which does not contain plasmid DNA; b) generating a library of chromosomal genomic DNA from said *Bacillus* species strain which does not contain plasmid DNA; c) obtaining the nucleotide

sequence of said chromosomal genomic DNA; d) identifying a *Bacillus* species strain which contains plasmid DNA; e) generating a library of said *Bacillus* species plasmid DNA; f) obtaining the nucleotide sequence of said plasmid DNA; g) subtracting any common sequences identified in the plasmid DNA which are also identified in the chromosomal genomic DNA; and h) constructing contigs and sequences of said plasmid DNA, wherein said contigs and sequences comprise the plasmid DNA sequence of said *Bacillus* species.

Insect inhibitory protein-encoding nucleic acids of the present invention will find particular uses in the plant protection against insects. For instance, insect-resistant transgenic plants can be generated by introducing the exogenous nucleic acids encoding an insect inhibitory polypeptide or protein or insect inhibitory fragment thereof listed in Table 1. Another example is to engineer transgenic microorganism (bacteria or fungi) to express insect inhibitory polypeptides or proteins of the present invention and then to apply them to the insect food source or allow them to reside in soil surrounding plant roots or on the surface of plant leaves.

The transgenic microorganisms of the present invention may be used to produce *B. thuringiensis* polypeptides or proteins of interest., particularly insect inhibitory polypeptides or proteins. Insect inhibitory polypeptides or proteins or insect inhibitory fragments thereof may be secreted, for example as in bacterial systems, meaning targeted to either the periplasm as for gram negative bacteria or localized to the extracellular space for gram negative or any other type of bacterium, or localized to the intracellular spaces within the cytoplasm. Such compositions may be administered to insects according to methods well known in the art. For example, insect inhibitory polypeptides or proteins of the present invention may be formulated as sprayable compositions or as a bait matrix.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

### Example 1

The DNA library designated LIB3237 (Monsanto Company, St. Louis, Missouri, United States of America) was prepared from purified chromosomal DNA of *Bacillus thuringiensis* strain EG10650. *B. thuringiensis* EG10650 is a derivative of strain EG10368 (U.S. Patent No. 5759538; June 2, 1998) that is deficient in neutral protease and alkaline protease activities and contains only one known extrachromosomal plasmid element of 7.5 kb. Deletion mutations in both the alkaline protease and neutral protease genes which were constructed first in strain EG10368 to produce strains EG10654 and EG10624 were combined to produce strain EG10650. Strain EG10650 provides the value described in

US 5,759,538 in which microorganisms containing these modifications exhibit markedly lower proteolytic activity with respect to counterpart microorganisms containing non-disabled protease genes. The result of the reduced proteolytic activity is that the microorganisms expressing insecticidal crystal protein genes and which contain such disabled protease genes produce higher levels of insecticidal crystal proteins and produce crystal proteins having increased stability during storage.

To assure the availability of materials to those interested members of the public upon issuance of a patent on the present application, deposits of the microorganisms listed above were made prior to filing the present application with the ARS Patent Collection, Agricultural Research Service Culture Collection (NRRL), 1815 North University Street, Peoria, Ill. 61604. These microorganism deposits were made under the provisions of the "Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure". EG10654 (NRRL Accession Number NRRL B-21344) and EG10624 (NRRL Accession Number NRRL B-21347) are currently available without restriction. All restrictions on the availability to the public of EG10650 (NRRL Accession Number NRRL B-30217) will be irrevocably removed upon issuance of a United States Patent based on this application. The DNA preparation, chromosomal DNA purification and library construction are described below.

#### DNA Preparation:

*Bacillus thuringiensis* EG10650 was grown under standard conditions in Brain Heart Infusion medium (Difco). Bacterial DNA was prepared in Agarose Plugs, according to the following procedure:

1. Single colony or loop full of bacteria was inoculated in 5 ml of plain Brain Heart Infusion medium and allowed to grow for overnight at 37 ° C;
2. One ml of above culture was inoculated in 20 ml of plain Brain Heart Infusion medium and allowed to grow till 0.6 OD<sub>600</sub>;
3. 180 µg/µl Chloramphenicol was added into the culture and the culture is allowed to grow for one more hour;
4. Cells were harvested by centrifuging at 5,000 g for 10 minutes at 4° C and the harvested cells were washed with 10 ml Solution A and then re-suspended in 1 ml of cold Solution A;
5. The re-suspended cells were brought to room temperature and quickly mixed with equal volume of 2% Seaplaque (FMC Corp.) agarose which was prepared in Solution A and kept at 50° C. The mixture of the cells and agarose was casted as 100 ul plugs in a mold placed on ice;

6. After 1 hour or solidification, the plugs were carefully removed from mold and placed in a tube containing 2 ml of Solution B. The plugs were incubated in a bath at 50°C for 12-16 hours with occasional shaking;
7. Proteinase K (0.5 mg/ml) was added in the tube and incubation was carried out for 16-24 hours; and
- 5 8. Finally the above solution was carefully replaced with 2 ml Solution C with 75 ug/ml Rnase. Incubation with Rnase was carried out for 16-24 hours. The plugs were stored at 4° C in the same solution untill they were needed.

Solution A: 1M NaCl, 50mMTris HCl (pH7.5)

Solution B: 50mMTris HCl (pH7.5), 100 mM EDTA, 100 mM NaCl and 2% lysozyme.

- 10 Solution C: 50mMTris HCl (pH7.5), 100 mM EDTA, 100 mM NaCl and 1% SDS.

#### Separation of Chromosomal DNA from Plasmid DNA:

- 15 High molecular weight plasmid DNA was removed from chromosomal DNA for obtaining desired DNA in plugs. Plasmids were removed By CHEF electrophoresis gel. Plugs were subjected to a electrical field of 6 V/cm at an angle of 120 degree in 0.5 X TBE buffer and 1% agarose gel. Initial switch time of 0.47 seconds and final switch time 44.69 seconds with linear ramping was used for 15 hours. After 15 hours plugs were retrieved form gel and were used for isolation of chromosomal DNA by standard methods.

#### Library Construction:

- 20 Purified chromosomal DNA of *Bacillus thuringiensis* strain EG 10650 was sheared to a desired size before size selecting by regular agarose gel electrophoresis. Ends of sheared and size selected DNA were repaired by PFU Polymerase (Stratagene). Repaired DNA was ligated to pSTBlue-1 vector at EcoR V site (Novagen) by using standard reagents and conditions recommended by manufacturer of DNA ligase enzyme (Life Technology). After ligation DNA was introduced by electroporation into
- 25 commercially available *E. coli* DH10B cells (Life Technology). Transformants were plated on LB agar plate with ampicillin for selection of plasmids. Plasmid DNA was prepared by standard procedure for sequencing by T7 and SP6 primers.

#### **Example 2**

- 30 Two basic methods can be used for DNA sequencing, the chain termination method of Sanger *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 74:5463-5467 (1977) and the chemical degradation method of Maxam and Gilbert, *Proc. Natl. Acad. Sci. (U.S.A.)* 74:560-564 (1977). Automation and advances in technology

such as the replacement of radioisotopes with fluorescence-based sequencing have reduced the effort required to sequence DNA (Craxton, *Method*, 2:20-26 (1991); Ju *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 92:4347-4351 (1995); Tabor and Richardson, *Proc. Natl. Acad. Sci. (U.S.A.)* 92:6339-6343 (1995)).

Automated sequencers are available from, for example, Pharmacia Biotech, Inc., Piscataway, New Jersey (Pharmacia ALF), LI-COR, Inc., Lincoln, Nebraska (LI-COR 4,000) and Millipore, Bedford, Massachusetts (Millipore BaseStation).

In addition, advances in capillary gel electrophoresis have also reduced the effort required to sequence DNA and such advances provide a rapid high resolution approach for sequencing DNA samples (Swerdlow and Gesteland, *Nucleic Acids Res.* 18:1415-1419 (1990); Smith, *Nature* 349:812-813 (1991); Luckey *et al.*, *Methods Enzymol.* 218:154-172 (1993); Lu *et al.*, *J. Chromatog. A.* 680:497-501 (1994); Carson *et al.*, *Anal. Chem.* 65:3219-3226 (1993); Huang *et al.*, *Anal. Chem.* 64:2149-2154 (1992); Kheterpal *et al.*, *Electrophoresis* 17:1852-1859 (1996); Quesada and Zhang, *Electrophoresis* 17:1841-1851 (1996); Baba, *Yakugaku Zasshi* 117:265-281 (1997)).

A number of sequencing techniques are known in the art, including fluorescence-based sequencing methodologies. These methods have the detection, automation and instrumentation capability necessary for the analysis of large volumes of sequence data. Currently, the 377 DNA Sequencer (Perkin-Elmer Corp., Applied Biosystems Div., Foster City, CA) allows the most rapid electrophoresis and data collection. With these types of automated systems, fluorescent dye-labeled sequence reaction products are detected and data entered directly into the computer, producing a chromatogram that is subsequently viewed, stored, and analyzed using the corresponding software programs. These methods are known to those of skill in the art and have been described and reviewed (Birren *et al.*, *Genome Analysis: Analyzing DNA*, 1, Cold Spring Harbor, New York).

### Example 3

#### Generation and Assembly of *Bacillus thuringiensis* Chromosomal Sequence

This example serves to illustrate the generation of the contigs and singletons listed in the Sequence Listing.

PHRED (phragment editor), which is developed by Phil Green at the University of Washington, was used to call the bases from the sequence trace files and to assign quality scores to the bases. PHRED uses Fourier methods to examine the four base traces in the region surrounding each point in the data set in order to predict a series of evenly spaced predicted locations. That is, it determines where the peaks would be centered if there were no compressions, dropouts, or other factors shifting the peaks from their "true" locations. Next, PHRED examines each trace to find the centers of the actual, or observed peaks



and the areas of these peaks relative to their neighbors. The peaks are detected independently along each of the four traces so many peaks overlap. A dynamic programming algorithm is used to match the observed peaks detected in the second step with the predicted peak locations found in the first step. Default parameters were used in the base calling.

After the base calling is completed, sequence preprocessing is performed. Quality assessment and trimming is performed by determining the maximum scoring segment of PHRED quality score  $> 10$ . Cloning sequences are removed by utilizing cross\_match (<http://www.mbt.washington.edu>) and searching a database of relevant cloning sequences. Contaminating sequences (E. coli, yeast, vector, linker) are then removed from the dataset by utilizing cross\_match to search a database of contamination sequences.

The preprocessed sequences are then assembled into contigs, or groups of overlapping sequences. Contigs are assembled using PHRAP (phragment assembly program), also developed by Green at the University of Washington (<http://www.mbt.washington.edu>) using default assembly parameters. This program takes a file of shotgun sequences and compiles consensus contig sequences. Alignments are influenced by quality scores, based on Green's algorithm. Singletons are the remaining sequences without sufficient overlaps with others after the assembly.

A total of 8283 contigs and singletons were obtained. Contig sequences are recognized as those sequences whose designations begin with . Singleton sequences are recognized as those having designations which begin with. All contig and singleton sequences were run through the annotation and gene selection processes as described in Example 4.

#### Example 4

##### Identification of *Bacillus thuringiensis* Genes

This example illustrates the identification of genes within the contig and singleton sequences assembled as described in Example 3. The genes and partial genes embedded in such contigs and singletons were identified through a series of informatic analyses.

Contigs and singletons are interrogated using AAT-NAP and BLASTP. AAT\_NAP is a program used for constructing a global alignment of a DNA sequence and a protein sequence (Huang, X. *et al.* Genomics 46:37-45 (1997), the entirety of which is herein incorporated by reference). The alignment model of NAP accommodates introns and frameshifts within codons. The scheme for scoring an alignment has several features that allow NAP to identify the exact locations of introns. A nucleotide insertion gap of length  $\leq k$  is given a linear penalty, and a nucleotide insertion gap of length  $> k$  is penalized as a gap of length  $k$ , where the value for  $k$  is the default value. The NAP program reports the starting and ending coordinates of predicted genes. The input to the NAP program includes the query

sequence, the protein database and a coordinate file produced by AAT\_EXT (an adapter between a database search program and a sequence alignment program) from the output of AAT\_DPS ( a program computing high-scoring chains of segment pairs between a query DNA sequence and the public non-redundant protein database from NCBI) The NAP program scans the protein database and finds the protein sequence for each coordinate record. Then for each coordinate record, NAP locates the query region, extends the region in both directions by a certain number of bases, and computes an alignment of the extended region and the protein sequence. NAP corrects frameshifts in the query sequence.

BLASTP is used to validate the amino acid sequences and hits reported by the AAT\_NAP program and to assign BLAST scores and p values to each sequence/hit pair. The AAT\_NAP generated amino acid sequences are compared with the public non-redundant protein database (nr.aa from NCBI) using the default BLASTP parameters except that the V parameter is set to 1000000 (to report up to 1000000 hits that exceed the BLASTP default report cutoff) If the hit reported by AAT\_NAP for a particular amino acid sequence is not reported by BLASTP, that particular amino acid sequence is removed. Protein encoding regions in the *Bacillus thuringiensis* nucleic acid molecules of the present invention are listed in Table 1.

Coding sequences identified in Table 1 encode many useful *B. thuringiensis* polypeptides or proteins or fragments thereof, including but not limited to insect inhibitory polypeptides or proteins, polypeptides or proteins capable of conferring antibiotic resistance, cytotoxin proteins which may be used as microbial inhibitory proteins including bactericidal, bacteriostatic, fungicidal, and fungistatic polypeptides or proteins, polyketide synthases, polypeptides or proteins capable of conferring resistance to heavy metals or other chemicals, transposons and mobile genetic elements and their corresponding transposases, excisases, integrases, and invertases, phage and phage particle proteins, transcription regulatory proteins, translation regulatory proteins, and other useful proteins homologous to proteins.

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1	Bt1Gc8697	Bt1G1	371-1	g3122850	196	238	2.10E-19	42	26	PREPROTEIN TRANSLOCASE SECA SUBUNIT [Bacillus firmus]
2	Bt1Gc8696	Bt1G2	1-422	g2497382	150	145	2.30E-09	34	33	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
3	Bt1Gc8693	Bt1G3	1-388	g1944414	112	165	4.30E-11	34	16	(D87026) glycogen phosphorylase [Bacillus stearothermophilus]
4	Bt1Gc8695	Bt1G4	387-1	g3122886	152	200	8.90E-15	38	15	ALANYL-TRNA SYNTHETASE (ALANINE--TRNA LIGASE) (ALARS) [Bacillus subtilis]
5	Bt1Gc4	Bt1G5	426-1	g585034	365	385	1.20E-35	55	32	D-ALANYL-D-ALANINE CARBOXYPEPTIDASE PRECURSOR (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) (CPASE) (PBPS) [Bacillus subtilis]
6	Bt1Gc3	Bt1G6	144-1	g2117582	173	175	6.40E-12	63	4	respiratory nitrate reductase alpha chain - Bacillus subtilis [Bacillus subtilis]
7	Bt1Gc6	Bt1G9	1-461	g4835715	58	140	2.70E-09	31	49	(AF064522) major autolysin [Moraxella sp. CK-1]
8	Bt1Gc12	Bt1G10	524-329	g2635805	211	227	6.70E-19	65	31	(Z99120) similar to two-component response regulator [Bacillus subtilis]
9	Bt1Gc10	Bt1G13	1-296	g2632778	140	143	9.30E-09	62	7	(Z99106) similar to hypothetical proteins [Bacillus subtilis]
9	Bt1Gc10	Bt1G14	1-296	g1107529	87	97	0.00022	35	22	(X88849) ceuC [Campylobacter coli]
10	Bt1Gc16	Bt1G15	615-1	g2127819	230	280	1.60E-24	32	73	cobalt transport ATP-binding protein O homolog - Methanococcus jannaschii [Methanococcus jannaschii]
11	Bt1Gc18	Bt1G16	260-1	g2494052	312	261	1.70E-22	70	31	PURINE NUCLEOSIDE PHOSPHORYLASE (INOSINE PHOSPHORYLASE) (PNP) [Bacillus stearothermophilus]
11	Bt1Gc18	Bt1G17	545-279	g2558482	354	363	2.60E-33	76	22	(AB008120) phosphopentomutase [Bacillus stearothermophilus]
12	Bt1Gc19	Bt1G18	496-1	g2635675	417	390	3.60E-36	53	44	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
13	Bt1Gc21	Bt1G20	1-501	g3123268	756	631	1.00E-61	87	67	30S RIBOSOMAL PROTEIN S2 (BS1) (VEGETATIVE PROTEIN 209) (VEG209) [Bacillus subtilis]
14	Bt1Gc20	Bt1G21	1-469	g1710383	584	606	4.60E-59	72	53	PROBABLE INTEGRASE/RECOMBINASE RIPX [Bacillus subtilis]
15	Bt1Gc22	Bt1G22	351-130	g586863	191	195	1.60E-15	50	99	HYPOTHETICAL 9.2 KD PROTEIN IN RECR-BOFA INTERGENIC REGION [Bacillus subtilis]
15	Bt1Gc22	Bt1G23	508-369	g585801	182	202	3.00E-16	85	23	RECOMBINATION PROTEIN RECR [Bacillus subtilis]
16	Bt1Gc27	Bt1G25	589-328	g2233359	125	206	5.70E-16	49	21	dehydratase, D-Ser [Escherichia coli]
17	Bt1Gc29	Bt1G26	370-1	g2634023	504	436	4.80E-41	82	51	(Z99112) uridylylate kinase [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
17	Bt1Gc29	Bt1G27	524-446	g3123214	87	108	1.20E-05	78	9	ELONGATION FACTOR TS (EF-TS) [Bacillus subtilis]
18	Bt1Gc34	Bt1G30	1-240	g2635587	111	87	0.00061	35	48	(Z99119) yuaE [Bacillus subtilis]
19	Bt1Gc32	Bt1G31	1-493	g3913682	69	84	0.0084	34	27	FORMAMIDASE (FORMAMIDE AMIDOHYDROLASE) [Methylophilus methylotrophus]
20	Bt1Gc30	Bt1G32	501-345	g1196398	203	91	0.00017	65	98	(M21305) unknown protein [Homo sapiens]
20	Bt1Gc30	Bt1G33	501-345	g1196398	203	91	0.00017	65	98	(M21305) unknown protein [Homo sapiens]
21	Bt1Gc36	Bt1G34	41-346	g1652388	176	114	9.80E-06	35	15	(D90905) transketolase [Synechocystis sp.]
21	Bt1Gc36	Bt1G35	1-346	g1620930	97	152	4.60E-10	42	16	(Z79580) putative orf [Bacillus subtilis]
22	Bt1Gc44	Bt1G38	1-551	g1945711	427	366	1.20E-33	47	44	(Z94043) hypothetical protein [Bacillus subtilis]
23	Bt1Gc45	Bt1G39	392-1	g3122013	689	627	2.70E-61	95	37	DNAJ PROTEIN [Streptococcus pneumoniae]
24	Bt1Gc43	Bt1G40	401-1	g2226238	517	522	3.70E-50	71	39	(Y14083) hypothetical protein [Bacillus subtilis]
25	Bt1Gc50	Bt1G41	478-1	g1171068	605	639	1.50E-62	77	37	PROBABLE UDP-N-ACETYLGLUCOSAMINE 1-CARBOXYVINYLTRANSFERASE (ENOYLPIRUVATE TRANSFERASE) (UDP-N-ACETYLGLUCOSAMINE ENOLPYRUVYL TRANSFERASE) (EPT) [Bacillus subtilis]
26	Bt1Gc48	Bt1G42	1-129	g2493270	126	144	2.30E-09	66	11	HYPOTHETICAL 41.2 KD PROTEIN CY277.09 [Mycobacterium tuberculosis]
27	Bt1Gc52	Bt1G43	1-502	g1730885	213	289	1.80E-25	35	46	HYPOTHETICAL 40.7 KD PROTEIN IN FER-RECQ INTERGENIC REGION [Bacillus subtilis]
28	Bt1Gc56	Bt1G44	1-511	g1934835	253	311	4.80E-27	56	16	(Z93940) asparagine synthetase [Bacillus subtilis]
29	Bt1Gc58	Bt1G46	1-326	g3334367	157	216	5.70E-17	48	22	HEMOLYSIN PRECURSOR [Bacillus cereus]
30	Bt1Gc59	Bt1G47	27-513	g2443255	481	510	6.90E-49	58	43	(D86417) YfmL [Bacillus subtilis]
31	Bt1Gc61	Bt1G49	128-563	g1731087	186	222	2.90E-18	36	42	HYPOTHETICAL 39.0 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
32	Bt1Gc63	Bt1G50	1-712	g118334	840	763	1.10E-75	68	48	LYSINE DECARBOXYLASE (LDC) [Bacillus subtilis]
33	Bt1Gc67	Bt1G51	1-618	g266515	250	273	2.10E-22	29	19	MULTIDRUG RESISTANCE-LIKE A TP-BINDING PROTEIN MDL [Escherichia coli]
34	Bt1Gc66	Bt1G53	458-1	g1770012	596	601	1.60E-58	72	42	(Z75208) hypothetical protein [Bacillus subtilis]
35	Bt1Gc68	Bt1G54	529-1	g2984251	97	178	6.10E-13	29	39	(AE000768) protoporphyrimogen oxidase [Aquifex aeolicus]
36	Bt1Gc65	Bt1G55	1-358	g1934609	305	235	1.90E-18	53	12	(U93874) formate dehydrogenase chain A [Bacillus subtilis]
37	Bt1Gc72	Bt1G57	1-333	g2688563	157	197	1.00E-15	37	41	(AE001165) spermidine/putrescine ABC transporter, permease protein (potB) [Borrelia burgdorferi]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
38	Bt1Gc73	Bt1G58	365-1	g1176701	425	156	3.20E-19	71	55	HYPOTHETICAL 23.6 KD PROTEIN IN QCRC-DAPB INTERGENIC REGION [Bacillus subtilis]
39	Bt1Gc76	Bt1G59	530-1	g3257185	459	482	6.40E-46	52	56	(AP000003) 314aa long hypothetical protein [Pyrococcus horikoshii]
40	Bt1Gc77	Bt1G60	1-463	g2635913	481	462	8.40E-44	60	50	(Z99121) similar to hypothetical proteins [Bacillus subtilis]
41	Bt1Gc75	Bt1G61	417-1	g1545847	239	263	2.40E-22	40	32	(U50396) WbpA [Pseudomonas aeruginosa]
42	Bt1Gc79	Bt1G62	59-505	g3582242	67	133	6.10E-09	33	87	(AE001272) L. lactis predicted coding region ORF00033 [Lactococcus lactis]
43	Bt1Gc82	Bt1G63	223-446	g2635847	248	267	1.20E-22	63	16	(Z99121) similar to ABC transporter (amino acid permease) [Bacillus subtilis]
44	Bt1Gc84	Bt1G64	1-494	g1075972	351	370	2.10E-33	45	25	tetB(P) protein - Clostridium perfringens [Clostridium perfringens]
45	Bt1Gc87	Bt1G65	1-299	g121563	274	291	1.10E-25	56	27	SPORE PROTEASE [Bacillus subtilis]
46	Bt1Gc88	Bt1G66	6-503	g4584100	389	399	4.00E-37	48	99	(AJ010132) hypothetical protein [Bacillus cereus]
47	Bt1Gc90	Bt1G68	1-290	g2909666	156	178	1.00E-13	40	45	(U68399) putative ABC transporter [Haemophilus influenzae]
48	Bt1Gc91	Bt1G69	1-377	g2495457	479	495	2.70E-47	77	32	HYPOTHETICAL 41.5 KD PROTEIN IN AMHX-AMYE INTERGENIC REGION [Bacillus subtilis]
49	Bt1Gc93	Bt1G70	1-441	g1673402	189	230	1.60E-18	33	30	(Z82044) hypothetical 54.4 kd protein [Bacillus subtilis]
50	Bt1Gc92	Bt1G72	1-987	g1805464	550	533	2.50E-51	40	61	(D50453) homologue of copper export protein PcoD of E. coli [Bacillus subtilis]
51	Bt1Gc95	Bt1G73	35-611	g3123166	282	306	2.90E-27	34	84	HYPOTHETICAL 25.7 KD PROTEIN IN GERAC-FHUC INTERGENIC REGION [Bacillus subtilis]
52	Bt1Gc97	Bt1G74	1-527	g4490571	89	146	2.60E-10	30	43	(AJ010302) mg protoporphyrin methyltransferase [Rhodobacter sphaeroides]
52	Bt1Gc97	Bt1G75	1-527	g4226092	104	178	3.90E-13	30	41	(AF125448) contains similarity to methyltransferases [Caenorhabditis elegans]
53	Bt1Gc98	Bt1G76	572-35	g2828523	511	487	1.90E-46	57	38	GABA PERMEASE (4-AMINO BUTYRATE TRANSPORT CARRIER) (GAMA-AMINO BUTYRATE PERMEASE) [Bacillus subtilis]
54	Bt1Gc99	Bt1G77	1-297	g2145962	292	250	2.40E-21	57	38	probable transport protein abc1 - Mycobacterium leprae [Mycobacterium leprae]
55	Bt1Gc101	Bt1G78	566-93	g2621742	102	159	1.10E-11	27	88	(AE000846) epoxidase [Methanobacterium thermoautotrophicum]
56	Bt1Gc100	Bt1G79	396-1	g2612883	234	281	1.30E-24	48	90	(AF015825) unknown [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
57	Bt1Gc103	Bt1G80	524-1	g2633967	604	643	5.50E-63	75	53	(Z99112) signal recognition particle (docking protein) [Bacillus subtilis]
58	Bt1Gc104	Bt1G81	114-553	g2621200	345	307	2.20E-27	46	85	(AE000804) ferritin like protein (RsgA) [Methanobacterium thermoautotrophicum]
59	Bt1Gc105	Bt1G82	1-488	g1172714	484	505	2.30E-48	60	50	PTS SYSTEM, GLUCOSE-SPECIFIC IIABC COMPONENT (EIIABC-GLC) (GLUCOSE-PERMEASE IIABC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, ABC COMPONENT) (EII-GLC / EIII-GLC) [Bacillus stearothermophilus]
60	Bt1Gc114	Bt1G84	1-234	g2148094	233	251	1.90E-21	51	23	TPP-dependent acetoin dehydrogenase alpha chain - Clostridium magnum [Clostridium magnum]
61	Bt1Gc112	Bt1G85	55-820	g133395	1123	1088	3.90E-110	88	81	DNA-DIRECTED RNA POLYMERASE ALPHA CHAIN (TRANSCRIPTASE ALPHA CHAIN) (RNA POLYMERASE ALPHA SUBUNIT) [Bacillus subtilis]
62	Bt1Gc110	Bt1G86	425-1	g1176967	111	151	3.00E-10	33	36	HYPOTHETICAL 37.6 KD PROTEIN IN GNTR-HTPG INTERGENIC REGION [Bacillus subtilis]
63	Bt1Gc115	Bt1G87	62-523	g418449	89	103	1.40E-05	22	81	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN HEMY-GLTT INTERGENIC REGION (ORFA) [Bacillus subtilis]
64	Bt1Gc111	Bt1G88	663-318	g3005554	93	140	4.30E-09	22	36	(AF047044) putative transposase [Anabaena PCC7120]
65	Bt1Gc116	Bt1G89	592-168	g2641974	204	201	3.80E-16	36	99	(AB005149) orf2 [Exiguobacterium acetylicum]
66	Bt1Gc118	Bt1G90	1-121	g227370	152	168	1.20E-12	83	44	DNA binding protein HU [Bacillus caldolyticus]
66	Bt1Gc118	Bt1G91	553-146	g1651880	228	243	1.40E-20	37	86	(D90900) hypothetical protein [Synechocystis sp.]
67	Bt1Gc119	Bt1G92	245-1	g3036880	234	252	1.70E-20	53	11	(AL022374) putative ATP-dependent DNA helicase [Streptomyces coelicolor]
68	Bt1Gc124	Bt1G94	549-1	g2791591	125	107	0.00019	23	41	(AL021287) hypothetical protein Rv2994 [Mycobacterium tuberculosis]
69	Bt1Gc123	Bt1G95	141-447	g137016	533	540	4.60E-52	100	99	HYPOTHETICAL PROTEIN UL125 [human herpesvirus 5]
70	Bt1Gc131	Bt1G96	843-1	g4539136	679	734	6.90E-72	50	24	(AL049478) putative DNA polymerase III alpha chain [Mycobacterium leprae]
71	Bt1Gc136	Bt1G97	1-687	g1561763	607	544	1.70E-52	51	34	(U67061) pullulanase [Bacteroides thetaiotaomicron]
72	Bt1Gc135	Bt1G98	68-1	g4584138	116	114	6.30E-07	96	52	(AJ010138) hypothetical protein [Bacillus cereus]
72	Bt1Gc135	Bt1G99	591-498	g417314	116	144	1.70E-09	91	9	MBL PROTEIN [Bacillus cereus]
73	Bt1Gc117	Bt1G100	1-731	g419952	125	135	1.30E-09	29	13	alpha-latroinsectotoxin precursor - black widow spider

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
74	Bt1Gc137	Bt1G101	290-1	g4468689	331	336	1.50E-29	68	13	(fragment) [Latrodectus tredecimguttatus] (AL035591) ABC excision nuclease subunit C [Streptomyces coelicolor]
75	Bt1Gc134	Bt1G102	1-242	g118716	242	188	3.20E-14	67	22	DNAJ PROTEIN [Bacillus subtilis]
75	Bt1Gc134	Bt1G103	274-501	g1730988	262	262	1.30E-22	66	24	PROBABLE METHYLTRANSFERASE [Bacillus subtilis]
76	Bt1Gc139	Bt1G104	120-525	g586809	398	425	7.00E-40	57	97	HYPOTHETICAL 15.6 KD PROTEIN IN PURA-DNAC INTERGENIC REGION [Bacillus subtilis]
77	Bt1Gc145	Bt1G105	1-635	g730600	785	808	1.80E-80	67	45	AMINO-ACID PERMEASE ROCC [Bacillus subtilis]
78	Bt1Gc153	Bt1G107	1-392	g3123299	480	495	2.70E-47	69	28	HYPOTHETICAL TRANSPORT PROTEIN IN NDHF-CSGA INTERGENIC REGION (ORF1) [Bacillus subtilis]
78	Bt1Gc153	Bt1G108	808-557	g4321580	79	114	3.60E-06	57	12	(AF050114) alginate lyase [Pseudomonas sp. W7]
79	Bt1Gc144	Bt1G109	1-595	g3915055	596	581	2.10E-56	56	75	EXTRAGENIC SUPPRESSOR PROTEIN SUHB HOMOLOG [Bacillus subtilis]
80	Bt1Gc155	Bt1G110	732-1	g2224770	961	898	5.30E-90	76	55	(Z97025) product similar to E. coli PhoH protein [Bacillus subtilis]
81	Bt1Gc157	Bt1G111	186-1	g1361340	92	100	0.00012	49	18	lin protein - Staphylococcus aureus [Staphylococcus aureus]
82	Bt1Gc159	Bt1G112	1-210	g132987	313	331	6.40E-30	89	39	50S RIBOSOMAL PROTEIN L5 [Bacillus subtilis]
83	Bt1Gc152	Bt1G113	590-1	g1903038	597	620	1.50E-60	60	73	(Z93102) hypothetical 30.6 kd protein [Bacillus subtilis]
84	Bt1Gc163	Bt1G114	317-416	g4584153	120	129	1.60E-08	79	97	(AJ010112) spore germination protein, GerPF [Bacillus cereus]
85	Bt1Gc165	Bt1G115	1-349	g1731085	255	257	4.40E-22	44	74	HYPOTHETICAL 17.9 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
86	Bt1Gc166	Bt1G116	48-636	g4584494	186	256	5.70E-22	34	88	(AL049587) hypothetical protein [Streptomyces coelicolor]
87	Bt1Gc164	Bt1G117	1-518	g606198	482	329	1.20E-29	50	35	(U18997) pantothenate permease [Escherichia coli]
88	Bt1Gc167	Bt1G118	595-130	g1001794	105	181	5.00E-14	26	92	(D64005) hypothetical protein [Synechocystis sp.]
89	Bt1Gc174	Bt1G120	1-290	g2226166	338	357	3.40E-32	65	17	(Y14080) hypothetical protein [Bacillus subtilis]
90	Bt1Gc171	Bt1G121	1-737	g3023410	478	457	2.80E-43	43	56	BRANCHED-CHAIN AMINO ACID TRANSPORT SYSTEM CARRIER PROTEIN (BRANCHED-CHAIN AMINO ACID UPTAKE CARRIER) [Bacillus subtilis]
91	Bt1Gc176	Bt1G122	509-1	g121562	671	681	5.20E-67	76	46	SPORE PROTEASE [Bacillus megaterium]
92	Bt1Gc177	Bt1G123	1-361	g3183562	527	543	2.20E-52	89	31	SUCCINYL-COA SYNTHETASE BETA CHAIN (SCS-ALPHA) (VEGETATIVE PROTEIN 63) (VEG63) [Bacillus subtilis]
93	Bt1Gc179	Bt1G124	1-468	g3183519	672	655	3.00E-64	85	36	GID PROTEIN [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
94	Bt1Gc180	Bt1G125	717-203	g129003	94	223	4.80E-18	25	42	6-AMINOHEXANOATE-DIMER HYDROLASE (NYLON OLIGOMERS DEGRADING ENZYME EII) [Flavobacterium sp.]
95	Bt1Gc183	Bt1G127	919-97	g4512401	1020	977	2.20E-98	73	39	(AB017508) fus homologue (identity of 87% to B. subtilis) [Bacillus halodurans]
96	Bt1Gc186	Bt1G128	1-367	g2500890	307	282	2.40E-24	55	26	AMINO ACID CARRIER PROTEIN ALST [Bacillus subtilis]
97	Bt1Gc187	Bt1G129	505-1	g266864	712	715	1.30E-70	80	39	ADENYLOSUCCINATE SYNTHETASE (IMP--ASPARTATE LIGASE) [Bacillus subtilis]
98	Bt1Gc189	Bt1G130	215-1	g3861147	104	114	4.80E-06	33	17	(AJ235272) BICYCLOMYCIN RESISTANCE PROTEIN (bcr1) [Rickettsia prowazekii]
99	Bt1Gc190	Bt1G131	385-619	g4467970	170	193	2.00E-14	43	15	(X76640) hypothetical protein [Myxococcus xanthus]
100	Bt1Gc191	Bt1G132	1-393	g2415743	434	395	1.10E-36	72	49	(AB000617) YceF [Bacillus subtilis]
101	Bt1Gc195	Bt1G133	1-543	g2293329	239	216	3.00E-17	32	46	(AF008220) YtdD [Bacillus subtilis]
102	Bt1Gc196	Bt1G134	1-147	g586206	178	178	1.90E-13	67	16	D-ALANYL-D-ALANINE CARBOXYPEPTIDASE (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) [Transposon Tn1546]
103	Bt1Gc199	Bt1G135	1-468	g1724006	493	266	1.50E-22	62	33	HYPOTHETICAL 49.0 KD PROTEIN IN CSPB-GLPP INTERGENIC REGION [Bacillus subtilis]
104	Bt1Gc197	Bt1G136	247-1	g1805394	310	276	9.60E-24	73	18	(D50453) homologue of proline permease of E. coli [Bacillus subtilis]
105	Bt1Gc201	Bt1G137	1-339	g1945654	129	181	5.00E-14	37	41	(Z94043) hypothetical protein [Bacillus subtilis]
106	Bt1Gc178	Bt1G138	1-477	g116230	654	625	4.50E-61	86	29	60 KD CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) [Bacillus subtilis]
107	Bt1Gc198	Bt1G139	1-684	g121879	95	167	1.50E-12	24	92	GLYCEROL OPERON REGULATORY PROTEIN [Streptomyces coelicolor]
108	Bt1Gc206	Bt1G140	1-148	g2226255	82	112	6.50E-06	49	13	(Y14084) hypothetical protein [Bacillus subtilis]
109	Bt1Gc210	Bt1G141	491-1	g2415742	601	554	1.50E-53	68	85	(AB000617) YceE [Bacillus subtilis]
110	Bt1Gc214	Bt1G142	173-1	g2739100	116	123	9.90E-07	41	9	(AF027507) primase [Mycobacterium smegmatis]
111	Bt1Gc215	Bt1G143	1-252	g2632528	375	375	1.40E-34	80	18	(Z99105) similar to sodium/proton-dependent alanine transporter [Bacillus subtilis]
112	Bt1Gc211	Bt1G144	1-525	g1165306	810	814	4.20E-81	89	63	(U43929) L2 [Bacillus subtilis]
113	Bt1Gc217	Bt1G145	1-137	g730959	97	121	1.70E-06	54	7	METHYL-ACCEPTING CHEMOTAXIS PROTEIN TLPB [Bacillus subtilis]
114	Bt1Gc218	Bt1G146	1-404	g135723	357	355	1.80E-32	52	41	THREONINE DEHYDRATASE CATABOLIC



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
115	Bt1Gc219	Bt1G147	1-138	g2495571	161	175	2.70E-13	62	16	(THREONINE DEAMINASE) [Escherichia coli]
116	Bt1Gc221	Bt1G148	409-1	g732321	373	363	2.60E-33	61	76	HYPOTHETICAL PROTEIN HI0105 [Haemophilus influenzae Rd]
117	Bt1Gc226	Bt1G149	500-1	g2815343	303	319	1.20E-28	35	46	HYPOTHETICAL 22.0 KD PROTEIN IN FLIT-SECA INTERGENIC REGION [Bacillus subtilis]
118	Bt1Gc225	Bt1G150	301-473	g1176995	128	156	3.00E-11	52	23	putative NTP pyrophosphohydrolase [Streptomyces coelicolor]
119	Bt1Gc223	Bt1G151	386-158	g1648861	220	141	6.40E-09	68	18	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN IN IDH 3'REGION [Bacillus subtilis]
120	Bt1Gc230	Bt1G152	562-1	g2632528	531	463	6.60E-44	53	40	(Z81356) UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Bacillus subtilis]
121	Bt1Gc232	Bt1G153	381-1	g1945676	171	194	2.10E-15	33	74	similar to sodium/proton-dependent alanine transporter [Bacillus subtilis]
122	Bt1Gc233	Bt1G154	1-682	g2632602	338	335	2.40E-30	32	100	hypothetical protein [Bacillus subtilis]
123	Bt1Gc237	Bt1G156	1-746	g1731017	584	617	3.20E-60	50	35	(Z99105) yegJ [Bacillus subtilis]
124	Bt1Gc238	Bt1G157	1-564	g2226166	514	473	5.70E-45	55	32	HYPOTHETICAL 80.1 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
125	Bt1Gc240	Bt1G158	1-453	g484466	72	147	4.10E-09	32	12	(Y14080) hypothetical protein [Bacillus subtilis]
126	Bt1Gc239	Bt1G159	1-574	g98515	896	912	1.70E-91	97	68	glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39) precursor - Bacillus circulans (strain IAM1165) [Bacillus circulans]
126	Bt1Gc239	Bt1G160	655-740	g2127278	129	145	3.30E-10	97	14	Bacillus thuringiensis [Bacillus thuringiensis]
127	Bt1Gc244	Bt1G161	1-754	g1881228	379	437	3.70E-41	34	55	flagellin B chain - Bacillus thuringiensis [Bacillus thuringiensis alesti]
128	Bt1Gc231	Bt1G162	1-669	g3025461	128	139	1.20E-08	23	49	flagellin chain A - Bacillus thuringiensis (subsp. alesti) [Bacillus thuringiensis alesti]
129	Bt1Gc251	Bt1G163	352-1	g1731061	470	478	1.70E-45	77	81	(AB001488) SIMILAR TO ENZYMES WHICH ACT VIA AN ATP-DEPENDENT COVALENT BINDING OF AMP TO THEIR SUBSTRATE. [Bacillus subtilis]
130	Bt1Gc252	Bt1G164	115-1	g585351	142	158	1.40E-11	77	17	(U58131) PhoR [Clostridium acetobutylicum]
131	Bt1Gc253	Bt1G165	1-507	g1684650	400	436	4.80E-41	50	36	HYPOTHETICAL 16.2 KD PROTEIN IN BMRU-ANSR INTERGENIC REGION [Bacillus subtilis]
132	Bt1Gc254	Bt1G166	1-564	g2635912	543	554	1.50E-53	52	65	CYTIDYLATE KINASE (CK) (CYTIDINE MONOPHOSPHATE KINASE) (CMP KINASE) [Bacillus subtilis]
										unknown [Bacillus subtilis]
										(Z82987) similar to transcriptional regulator (LysR family) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
133	Bt1Gc257	Bt1G167	609-1	g1711354	196	268	7.60E-23	47	26	[Bacillus subtilis] PROBABLE D-SERINE DEHYDRATASE (D-SERINE DEAMINASE) [Bacillus subtilis]
134	Bt1Gc255	Bt1G168	189-839	g1405445	727	731	2.60E-72	64	100	(Z73234) YneB [Bacillus subtilis]
135	Bt1Gc256	Bt1G169	1-574	g2621572	52	140	1.20E-08	28	34	(AE000833) 5-methylcytosine-specific restriction enzyme McrB related protein [Methanobacterium thermoautotrophicum]
136	Bt1Gc258	Bt1G170	41-583	g2293312	661	634	5.00E-62	73	45	(AF008220) YtrP [Bacillus subtilis]
137	Bt1Gc259	Bt1G172	1-536	g2635181	723	658	1.40E-64	72	58	(Z99117) similar to protease [Bacillus subtilis]
138	Bt1Gc261	Bt1G173	1-602	g2632528	653	664	3.30E-65	63	42	(Z99105) similar to sodium/proton-dependent alanine transporter [Bacillus subtilis]
139	Bt1Gc263	Bt1G174	290-1	g1763706	185	155	2.90E-11	39	94	(Z83337) ywpE [Bacillus subtilis]
140	Bt1Gc268	Bt1G175	821-1	g1669857	212	154	2.10E-10	25	66	(U70055) macrolide-efflux protein [Streptococcus pyogenes]
141	Bt1Gc272	Bt1G176	242-1	g2135326	68	96	0.0011	33	10	H-NUC - human [Homo sapiens]
142	Bt1Gc270	Bt1G177	477-1	g730437	561	416	6.30E-39	72	37	URACIL PERMEASE (URACIL TRANSPORTER) [Bacillus caldolyticus]
143	Bt1Gc276	Bt1G179	1-562	g1110441	157	141	2.60E-09	29	65	(M80628) hypothetical product [Streptomyces griseus]
144	Bt1Gc279	Bt1G180	397-65	g1730957	321	365	1.60E-33	71	99	HYPOTHETICAL 11.6 KD PROTEIN IN COTD-KDUD INTERGENIC REGION [Bacillus subtilis]
145	Bt1Gc280	Bt1G181	513-1	g1174516	499	526	1.30E-49	55	17	ISOLEUCYL-TRNA SYNTHETASE, MUPIROCIN RESISTANT (ISOLEUCINE--TRNA LIGASE) (ILERS) (MUPIROCIN RESISTANCE PROTEIN) [Staphylococcus aureus]
146	Bt1Gc283	Bt1G182	453-1	g1842440	454	454	5.90E-43	57	36	(U87792) ClnA [Bacillus subtilis]
147	Bt1Gc286	Bt1G184	1-287	g2633910	218	240	2.80E-20	47	35	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
147	Bt1Gc286	Bt1G185	296-703	g3183467	300	352	3.80E-32	49	59	HYPOTHETICAL 25.7 KD PROTEIN IN SIGG-DIVIVA INTERGENIC REGION [Bacillus subtilis]
148	Bt1Gc284	Bt1G186	1-467	g606198	276	285	1.40E-24	41	33	(U18997) pantothenate permease [Escherichia coli]
149	Bt1Gc288	Bt1G187	838-63	g1842438	336	260	2.10E-22	32	85	(U87792) unknown [Bacillus subtilis]
150	Bt1Gc289	Bt1G188	683-159	g3688416	603	509	8.80E-49	62	81	(AJ006274) acetyltransferase-like protein [Lactobacillus sakei]
151	Bt1Gc290	Bt1G189	1-342	g1730267	88	158	1.40E-11	39	42	HYPOTHETICAL 20.5 KD PROTEIN IN HMP 3'REGION [Bacillus subtilis]
152	Bt1Gc292	Bt1G190	1-632	g1706650	824	873	2.40E-87	79	70	PROBABLE ENDONUCLEASE IV

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
153	Bt1Gc285	Bt1G191	620-120	g4894313	215	282	1.00E-24	37	99	(ENDODEOXYRIBONUCLEASE IV) [Bacillus subtilis]
154	Bt1Gc293	Bt1G192	755-190	g2407933	209	220	3.70E-18	28	61	(AF065404) pXO1-97 [Bacillus anthracis]
155	Bt1Gc295	Bt1G193	254-595	g2073397	223	195	1.60E-15	44	98	(AJ000325) putative membrane protein [Lactococcus lactis]
156	Bt1Gc297	Bt1G194	910-1	g3821797	815	835	2.50E-83	52	58	(Y09322) membrane bound protein LytR [Bacillus cereus]
157	Bt1Gc294	Bt1G195	957-297	g3602952	233	294	5.30E-26	31	94	(D10594) chitinase D precursor [Bacillus circulans]
158	Bt1Gc298	Bt1G196	427-1	g2635763	513	518	9.80E-50	74	31	(AF041468) ompR homolog [Guillardia theta]
159	Bt1Gc299	Bt1G197	530-799	g481591	94	129	1.60E-08	28	47	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
160	Bt1Gc296	Bt1G198	16-552	g2635183	297	290	1.40E-25	38	50	hypothetical protein 4 - Clostridium pasteurianum [Clostridium pasteurianum]
161	Bt1Gc300	Bt1G199	308-1	g2340010	211	224	1.40E-18	43	59	(Z99117) similar to folate metabolism [Bacillus subtilis]
161	Bt1Gc300	Bt1G200	462-553	g2340009	93	111	1.00E-05	65	7	(Z98682) YlbN protein [Bacillus subtilis]
162	Bt1Gc301	Bt1G201	1-659	g1684651	432	314	4.00E-28	37	56	(Z98682) YlbM protein [Bacillus subtilis]
163	Bt1Gc302	Bt1G202	479-694	g2984723	300	325	2.80E-29	89	99	(Z82987) unknown similar to quinolon resistance protein NorA [Bacillus subtilis]
164	Bt1Gc303	Bt1G203	451-1	g732327	492	491	7.10E-47	60	48	(AF053927) probable spore germination protein F [Bacillus cereus]
165	Bt1Gc305	Bt1G204	125-670	g3123238	589	628	2.20E-61	75	48	PROBABLE 1,4-DIHYDROXY-2-NAPHTHOATE OCTAPRENYLTRANSFERASE (DHNA-OCTAPRENYLTRANSFERASE) [Bacillus subtilis]
166	Bt1Gc308	Bt1G205	1-456	g1730889	260	270	1.90E-23	39	58	PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT (S COMPLEX, 42 KD SUBUNIT) (VEGETATIVE PROTEIN 220) (VEG220) [Bacillus subtilis]
166	Bt1Gc308	Bt1G206	471-655	g3025068	100	128	4.30E-08	37	25	HYPOTHETICAL 28.6 KD PROTEIN IN RECQ-CMK INTERGENIC REGION PRECURSOR [Bacillus subtilis]
167	Bt1Gc310	Bt1G207	1-216	g2632024	159	167	1.50E-12	49	52	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN ICDC-MINE INTERGENIC REGION [Escherichia coli]
167	Bt1Gc310	Bt1G208	1-1128	g4240001	168	271	4.70E-23	34	36	(AJ002571) YkhA [Bacillus subtilis]
168	Bt1Gc311	Bt1G209	1-726	g2632987	891	898	5.30E-90	70	53	(AB017186) cardiolipin synthase [Clostridium perfringens]
169	Bt1Gc312	Bt1G210	1-324	g4104605	229	263	1.00E-22	46	47	(Z99107) alternate gene name: yerS; similar to RNA methyltransferase [Bacillus subtilis]
169	Bt1Gc312	Bt1G211	402-525	g225559	187	203	1.40E-15	98	9	putative response regulator [Lactobacillus sakei]
170	Bt1Gc313	Bt1G212	1-809	g113811	1120	1070	3.10E-108	72	52	(AF036967) ORF IS231C [Bacillus thuringiensis]
										ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
171	Bt1Gc309	Bt1G213	1159-1	g1075972	806	880	4.30E-88	46	59	GLUCAN GLUCANOHYDROLASE) [Bacillus amyloquifaciens]
172	Bt1Gc316	Bt1G214	721-1	g2635792	235	297	7.50E-26	28	47	tet(P) protein - Clostridium perfringens [Clostridium perfringens]
173	Bt1Gc315	Bt1G215	191-1	g1730911	212	228	1.10E-18	69	17	(Z99120) similar to oligoendopeptidase [Bacillus subtilis]
174	Bt1Gc319	Bt1G216	1-506	g1706795	374	332	5.00E-30	45	43	HYPOTHETICAL 42.6 KD PROTEIN IN BSAA-IL VD INTERGENIC REGION [Bacillus subtilis]
175	Bt1Gc317	Bt1G217	787-1	g2127819	261	363	2.60E-33	31	94	FERRICHRONE TRANSPORT PERMEASE PROTEIN FHUB [Bacillus subtilis]
176	Bt1Gc322	Bt1G218	424-1	g2618995	333	336	1.90E-30	46	85	cobalt transport ATP-binding protein O homolog - Methanococcus jannaschii [Methanococcus jannaschii]
177	Bt1Gc321	Bt1G219	24-762	g466778	841	698	8.20E-69	63	50	putative alanine acetyl transferase [Bacillus subtilis]
178	Bt1Gc324	Bt1G220	1-318	g415453	285	257	1.30E-21	48	23	[Helicobacter pylori J99]
179	Bt1Gc327	Bt1G221	744-542	g2649585	124	179	1.00E-13	45	24	(AE001034) ABC transporter, ATP-binding protein [Archaeoglobus fulgidus]
180	Bt1Gc330	Bt1G222	1-243	g1945096	183	192	1.70E-14	54	19	lysine specific permease [Escherichia coli]
181	Bt1Gc328	Bt1G223	548-1	g96807	509	561	1.20E-53	58	20	putative SODIUM/ALANINE SYMPORTER [P31141 (492) transmembrane [Bacillus subtilis]
182	Bt1Gc329	Bt1G224	1-428	g2633471	539	560	3.50E-54	73	34	Mg2+-transporting ATPase (EC 3.6.1.-) mgfB - Salmonella typhimurium [Salmonella typhimurium]
183	Bt1Gc332	Bt1G225	609-1	g2499210	684	698	8.20E-69	67	97	(Z99109) similar to 3-oxoacyl- acyl-carrier protein synthase [Bacillus subtilis]
183	Bt1Gc332	Bt1G226	922-634	g1881373	227	245	8.30E-21	49	59	PUTATIVE NAD(P)H NITROREDUCTASE YDGI [Bacillus subtilis]
184	Bt1Gc333	Bt1G227	1-381	g548832	114	113	1.80E-13	29	76	(AB001488) SIMILAR TO SALMONELLA TYPHIMURIUM SLY GENE REQUIRED FOR SURVIVAL IN MACROPHAGE. [Bacillus subtilis]
185	Bt1Gc334	Bt1G228	1-549	g2633168	446	493	4.40E-47	51	58	RNA POLYMERASE SIGMA-H FACTOR (SIGMA-30) [Pseudomonas aeruginosa]
186	Bt1Gc335	Bt1G229	1033-782	g225559	409	409	3.50E-38	98	17	(Z99108) similar to iron(III) dicitrate transport permease [Bacillus subtilis]
187	Bt1Gc336	Bt1G230	1-400	g549496	154	200	4.90E-16	34	65	ORF IS231C [Bacillus thuringiensis]
										HYPOTHETICAL 22.1 KD PROTEIN IN HEML-PFS

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
188	Bt1Gc337	Bt1G231	1-824	g1731076	695	723	1.80E-71	50	50	INTERGENIC REGION [Escherichia coli]
189	Bt1Gc339	Bt1G232	548-965	g682770	417	417	4.90E-39	60	41	HYPOTHETICAL 61.8 KD PROTEIN IN GLNQ-ANSR [Bacillus subtilis]
190	Bt1Gc340	Bt1G233	60-313	g144312	125	143	5.30E-10	38	68	INTERGENIC REGION [Escherichia coli]
191	Bt1Gc341	Bt1G234	813-82	g140635	610	580	2.60E-56	47	91	(X57583) mceF [Escherichia coli]
192	Bt1Gc342	Bt1G235	1-410	g3023262	461	498	1.30E-47	67	35	(J01566) 13.8 kd ORF [Plasmid Cole1]
193	Bt1Gc347	Bt1G236	1-394	g2634307	424	334	3.10E-30	64	74	HYPOTHETICAL 28.6 KD PROTEIN IN FABB-MEPA [Escherichia coli]
193	Bt1Gc347	Bt1G237	517-706	g3023263	208	224	3.60E-18	67	16	INTERGENIC REGION [Escherichia coli]
194	Bt1Gc350	Bt1G238	684-1	g1507711	484	481	8.10E-46	46	40	PROBABLE NADH-DEPENDENT BUTANOL DEHYDROGENASE 1 [Bacillus subtilis]
195	Bt1Gc351	Bt1G239	431-1	g1661219	366	423	1.10E-39	52	28	(Z99114) similar to hypothetical proteins [Bacillus subtilis]
196	Bt1Gc352	Bt1G240	674-285	g2897873	285	287	2.90E-25	44	60	PROBABLE NADH-DEPENDENT BUTANOL DEHYDROGENASE 2 [Bacillus subtilis]
197	Bt1Gc320	Bt1G241	711-168	g732998	307	348	1.00E-31	36	99	(L80006) indolepyruvate decarboxylase [Erwinia herbicola]
198	Bt1Gc353	Bt1G242	645-1	g1881359	111	167	1.50E-12	24	94	(U75904) ProP [Escherichia coli]
199	Bt1Gc356	Bt1G243	584-273	g585113	140	175	2.20E-13	54	23	(U81516) unknown [Myxococcus xanthus]
200	Bt1Gc349	Bt1G244	1-554	g2293447	544	550	4.00E-53	57	48	(U01945) streptothricine-acetyl-transferase [Campylobacter coli]
201	Bt1Gc357	Bt1G245	566-130	g2633223	335	312	6.60E-28	50	65	(AB001488) SIMILAR TO YDFS GENE PRODUCT OF THIS ENTRY (YDFS_BACSU). [Bacillus subtilis]
202	Bt1Gc358	Bt1G246	1-418	g2443255	192	123	4.40E-07	43	36	EXODEOXYRIBONUCLEASE [Bacillus subtilis]
203	Bt1Gc362	Bt1G247	1-697	g2119798	268	396	8.30E-37	39	64	(AF008930) ATPase [Bacillus subtilis]
204	Bt1Gc364	Bt1G248	179-1	g3308998	138	135	3.80E-09	53	29	(Z99108) yhbJ [Bacillus subtilis]
204	Bt1Gc364	Bt1G249	959-432	g136472	237	284	6.10E-25	33	99	(D86417) YfmL [Bacillus subtilis]
205	Bt1Gc365	Bt1G250	413-1	g3915864	67	127	9.90E-08	32	28	penicillin binding protein 5 - Bacillus subtilis (fragment) [Bacillus subtilis]
206	Bt1Gc368	Bt1G251	1-397	g1730977	119	144	4.20E-10	27	63	(AB015998) CBP21 precursor [Serratia marcescens]
206	Bt1Gc368	Bt1G252	576-447	g2239276	178	147	2.20E-09	93	8	ACETYLTRANSFERASE (TABTOXIN RESISTANCE PROTEIN) [Pseudomonas syringae]
207	Bt1Gc367	Bt1G253	10-210	g2493761	348	348	1.00E-31	99	99	SMF PROTEIN [Bacillus subtilis]
208	Bt1Gc366	Bt1G254	605-442	g1405460	104	138	1.80E-09	57	36	HYPOTHETICAL 22.9 KD PROTEIN IN NUCB-AROD INTERGENIC REGION [Bacillus subtilis]
										(L77763) neutral protease A [Bacillus thuringiensis]
										MAJOR COLD SHOCK PROTEIN CSPA [Bacillus cereus]
										(Z73234) YneT [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
209	Bt1Gc370	Bt1G256	576-92	g2635763	705	728	5.50E-72	88	35	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
210	Bt1Gc372	Bt1G257	1-205	g3880839	139	159	2.60E-11	48	23	(AL031633) predicted using Genefinder; similar to Alcohol/other dehydrogenases, short chain type, Short chain dehydrogenase/reductase C-terminus; cDNA EST EMBL: C08550 comes from this gene [Caenorhabditis elegans]
211	Bt1Gc374	Bt1G258	1-633	g4033506	714	754	9.60E-75	70	73	FRUCTOSE-BISPHOSPHATE ALDOLASE [Bacillus stearothermophilus]
212	Bt1Gc376	Bt1G259	1-126	g1945051	152	153	2.50E-10	67	11	(U63928) L1 protein [Bacillus cereus]
212	Bt1Gc376	Bt1G260	239-567	g2507017	330	263	1.00E-22	68	29	HEMOLYSIN BL BINDING COMPONENT PRECURSOR (ENTEROTOXIN 40 KD SUBUNIT) [Bacillus cereus]
213	Bt1Gc377	Bt1G261	1-646	g3256693	139	214	9.30E-17	26	45	(AP000001) 484aa long hypothetical protein [Pyrococcus horikoshii]
214	Bt1Gc380	Bt1G262	463-1	g2633757	231	288	2.30E-25	50	31	(Z99111) similar to Xaa-Pro dipeptidase [Bacillus subtilis]
215	Bt1Gc379	Bt1G263	1-589	g3123227	693	744	1.10E-73	69	39	GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE) [Bacillus subtilis]
216	Bt1Gc373	Bt1G264	1-721	g2226222	309	369	6.00E-34	42	65	(Y14082) hypothetical protein [Bacillus subtilis]
217	Bt1Gc382	Bt1G265	1-1117	g1169000	1178	1181	5.40E-120	57	34	MICROBIAL COLLAGENASE PRECURSOR (120 KD COLLAGENASE) [Clostridium perfringens]
218	Bt1Gc383	Bt1G266	791-27	g3688826	336	277	4.50E-24	30	61	(AF084104) NatB [Bacillus firmus]
219	Bt1Gc388	Bt1G267	642-1	g2293198	445	443	8.70E-42	42	40	(AF008220) YtgP [Bacillus subtilis]
220	Bt1Gc387	Bt1G268	671-1	g2501678	931	857	1.20E-85	78	74	HYPOTHETICAL 33.7 KD PROTEIN IN LSP-PYRR INTERGENIC REGION (ORF-X) []
221	Bt1Gc384	Bt1G269	1-510	g1706797	308	355	1.80E-32	42	51	FERRICHRONE TRANSPORT PERMEASE PROTEIN FHUG [Bacillus subtilis]
222	Bt1Gc386	Bt1G270	570-1	g2274944	575	469	1.50E-44	58	48	(AJ000346) NapC protein [Enterococcus hirae]
223	Bt1Gc390	Bt1G271	1-578	g1934809	461	409	3.50E-38	47	36	(Z93937) unknown [Bacillus subtilis]
224	Bt1Gc392	Bt1G272	679-1	g1708643	456	500	7.90E-48	43	35	(U51115) YebA [Bacillus subtilis]
225	Bt1Gc393	Bt1G273	174-97	g586023	118	118	2.40E-07	85	96	STAGE V SPOULATION PROTEIN M [Bacillus subtilis]
225	Bt1Gc393	Bt1G274	879-244	g2337809	492	516	1.60E-49	47	100	(Y13937) YloS protein [Bacillus subtilis]
226	Bt1Gc396	Bt1G275	151-697	g4877899	226	301	9.70E-27	46	41	(AF138876) unknown [Bacillus thuringiensis]
226	Bt1Gc396	Bt1G276	432-697	g549114	374	410	2.70E-38	80	19	TRANSPPOSASE FOR INSERTION SEQUENCE ELEMENT IS231F [Bacillus thuringiensis]
227	Bt1Gc398	Bt1G277	1-453	g3183519	660	609	2.20E-59	84	35	GID PROTEIN [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
228	Bt1Gc400	Bt1G278	1129-176	g2293322	638	607	3.60E-59	44	71	(AF008220) branch-chain amino acid transporter [Bacillus subtilis]
229	Bt1Gc402	Bt1G279	227-1	g1173414	169	151	9.40E-10	49	13	PROTEIN-EXPORT MEMBRANE PROTEIN SEC2 [Haemophilus influenzae Rd]
230	Bt1Gc404	Bt1G280	759-240	g732390	645	652	6.20E-64	72	55	DEOXYRIBONUCLEOSIDE REGULATOR [Bacillus subtilis]
231	Bt1Gc405	Bt1G281	1-654	g2635801	764	716	1.00E-70	66	47	(Z99120) fumarate hydratase [Bacillus subtilis]
232	Bt1Gc395	Bt1G282	608-237	g1730948	180	99	2.50E-05	38	99	HYPOTHETICAL 14.5 KD PROTEIN IN PONA-COTD INTERGENIC REGION [Bacillus subtilis]
233	Bt1Gc408	Bt1G283	1-247	g1730905	154	165	2.50E-12	40	68	HYPOTHETICAL 13.7 KD PROTEIN IN BCSA-DEGR INTERGENIC REGION [Bacillus subtilis]
233	Bt1Gc408	Bt1G284	242-421	g1730906	164	169	9.40E-13	48	98	HYPOTHETICAL 7.2 KD PROTEIN IN BCSA-DEGR INTERGENIC REGION [Bacillus subtilis]
234	Bt1Gc410	Bt1G285	1-757	g2499158	828	862	3.40E-86	66	81	HYPOTHETICAL 34.8 KD PROTEIN IN HLYA 3'REGION [Edwardsiella tarda]
234	Bt1Gc410	Bt1G286	898-1415	g4894226	238	317	1.90E-28	59	31	(AF065404) pXO1-10 [Bacillus anthracis]
235	Bt1Gc406	Bt1G287	809-352	g1705506	493	511	5.40E-49	60	96	GLUTATHIONE PEROXIDASE HOMOLOG BSAA [Bacillus subtilis]
236	Bt1Gc411	Bt1G288	1-284	g3445480	373	384	1.50E-35	69	18	(AB008763) amylase [Bacillus sp.]
237	Bt1Gc409	Bt1G289	494-1	g118797	306	324	3.50E-29	38	44	DNA POLYMERASE III, BETA CHAIN [Bacillus subtilis]
238	Bt1Gc415	Bt1G290	250-608	g1176478	147	214	1.30E-16	36	22	ABC TRANSPORTER ATP-BINDING PROTEIN HI1252 [Haemophilus influenzae Rd]
239	Bt1Gc412	Bt1G291	1-511	g585132	261	300	1.20E-26	36	55	FERRICHROME-BINDING PROTEIN PRECURSOR [Bacillus subtilis]
239	Bt1Gc412	Bt1G292	459-1026	g1706795	177	201	1.30E-15	33	50	FERRICHROME TRANSPORT PERMEASE PROTEIN FHUB [Bacillus subtilis]
240	Bt1Gc417	Bt1G293	613-1	g2635863	575	560	4.50E-54	53	25	(Z99121) similar to heavy metal-transporting ATPase [Bacillus subtilis]
241	Bt1Gc418	Bt1G294	1-312	g1944618	225	238	4.60E-20	45	36	(Y12602) acid phosphatase [Streptococcus equisimilis]
242	Bt1Gc420	Bt1G295	102-1	g3023407	133	133	4.50E-08	74	8	BRANCHED-CHAIN AMINO ACID TRANSPORT SYSTEM CARRIER PROTEIN (BRANCHED-CHAIN AMINO ACID UPTAKE CARRIER) [Corynebacterium glutamicum]
243	Bt1Gc422	Bt1G296	1-294	g1731127	337	337	1.50E-30	64	33	HYPOTHETICAL 32.9 KD PROTEIN IN BLTR-SPOIIIC INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
243	Bt1Gc422	Bt1G297	329-965	g1731125	325	369	6.00E-34	43	99	HYPOTHETICAL 20.7 KD PROTEIN IN BL-TR-SPOIIC INTERGENIC REGION [Bacillus subtilis]
244	Bt1Gc423	Bt1G298	937-507	g1575577	238	265	6.30E-23	38	62	(U67196) DNA-binding response regulator [Thermotoga maritima]
245	Bt1Gc425	Bt1G299	564-1	g4580028	202	299	2.50E-26	37	35	(AF086638) CumA precursor [Pseudomonas putida GB-1]
246	Bt1Gc429	Bt1G300	338-1	g586703	221	129	1.10E-07	40	28	HYPOTHETICAL 43.0 KD PROTEIN IN PROK-TAG INTERGENIC REGION (F402) [Escherichia coli]
247	Bt1Gc428	Bt1G301	523-1	g2633180	96	127	2.60E-08	31	76	(Z99108) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
248	Bt1Gc430	Bt1G302	188-472	g628925	290	240	2.90E-19	61	14	phosphotransferase system enzyme II (EC 2.7.1.69), glucose-specific, factor IIA - Staphylococcus carnosus [Staphylococcus carnosus]
249	Bt1Gc431	Bt1G303	625-1	g2635863	809	747	5.30E-74	76	26	(Z99121) similar to heavy metal-transporting ATPase [Bacillus subtilis]
250	Bt1Gc434	Bt1G304	1-545	g2634067	307	344	2.70E-31	36	46	(Z99112) penicillin-binding protein [Bacillus subtilis]
251	Bt1Gc435	Bt1G305	485-1	g2126822	337	402	1.90E-37	51	61	bacteriolytic enzyme lytP - Bacillus sp [Bacillus sp.]
252	Bt1Gc438	Bt1G306	489-28	g586897	489	468	1.90E-44	60	99	TRANSCRIPTIONAL REGULATOR CTSR [Bacillus subtilis]
253	Bt1Gc439	Bt1G307	1-733	g2829802	263	306	6.60E-27	34	47	HYPOTHETICAL 53.3 KD PROTEIN IN SFP-GERKA INTERGENIC REGION [Bacillus subtilis]
254	Bt1Gc436	Bt1G308	1-136	g2916905	94	115	5.20E-06	39	9	(AL022004) hypothetical protein Rv0846c [Mycobacterium tuberculosis]
255	Bt1Gc437	Bt1G309	314-920	g1177018	745	675	2.30E-66	75	44	HYPOTHETICAL 48.3 KD PROTEIN IN KATB 3'REGION [Bacillus subtilis]
256	Bt1Gc442	Bt1G310	145-1051	g2415736	512	447	3.30E-42	38	94	(AB000617) YcdH [Bacillus subtilis]
257	Bt1Gc441	Bt1G311	681-1	g586875	933	933	1.00E-93	76	89	HYPOTHETICAL 29.2 KD PROTEIN IN METS-KSGA INTERGENIC REGION [Bacillus subtilis]
258	Bt1Gc444	Bt1G312	1-816	g3261741	357	423	1.10E-39	31	57	(Z95120) hypothetical protein Rv3225c [Mycobacterium tuberculosis]
259	Bt1Gc453	Bt1G313	1-919	g732351	406	419	3.00E-39	35	79	HYPOTHETICAL 43.3 KD PROTEIN IN QOXD-VPR INTERGENIC REGION [Bacillus subtilis]
260	Bt1Gc450	Bt1G314	515-1	g2618854	442	442	1.10E-41	48	70	(AF017113) YvoA [Bacillus subtilis]
260	Bt1Gc450	Bt1G315	918-529	g3122426	238	293	6.80E-26	47	53	GLUCOSAMINE-6-PHOSPHATE ISOMERASE (GLUCOSAMINE-6-PHOSPHATE DEAMINASE) [Bacillus subtilis]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
261	Bt1Gc443	Bt1G316	600-238	g2622920	317	357	1.10E-32	50	36	(AE000933) dTDP-glucose 4,6-dehydratase [Methanobacterium thermoautotrophicum]
262	Bt1Gc451	Bt1G317	152-643	g1945654	251	308	1.70E-27	42	65	(Z94043) hypothetical protein [Bacillus subtilis]
263	Bt1Gc449	Bt1G318	145-510	g4584112	197	273	9.00E-24	39	37	(AJ007510) 5-methylcytosine-specific restriction enzyme [Bacillus cereus]
264	Bt1Gc452	Bt1G319	751-262	g1711354	491	528	8.50E-51	64	36	PROBABLE D-SERINE DEHYDRATASE (D-SERINE DEAMINASE) [Bacillus subtilis]
265	Bt1Gc446	Bt1G320	1-335	g732151	308	319	1.20E-28	53	93	HYPOTHETICAL 14.0 KD PROTEIN IN LYSA 3'REGION (ORF3) [Bacillus methanolicus]
265	Bt1Gc446	Bt1G321	660-362	g141355	144	160	8.40E-12	30	57	HYPOTHETICAL 21.0 KD PROTEIN IN RIBT-DACB INTERGENIC REGION (ORFX6) [Bacillus subtilis]
266	Bt1Gc458	Bt1G322	1-890	g118704	1219	1178	1.10E-119	79	66	CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA [Bacillus subtilis]
267	Bt1Gc459	Bt1G323	1-565	g3183562	806	687	1.20E-67	83	49	SUCCINYL-COA SYNTHETASE BETA CHAIN (SCS-ALPHA) (VEGETATIVE PROTEIN 63) (VEG63) [Bacillus subtilis]
268	Bt1Gc460	Bt1G324	917-173	g1934789	154	178	1.90E-21	23	100	(Z93934) unknown [Bacillus subtilis]
269	Bt1Gc455	Bt1G326	786-1	g1731006	860	878	6.90E-88	63	59	PROBABLE RNA HELICASE IN CCCA-SODA INTERGENIC REGION [Bacillus subtilis]
270	Bt1Gc469	Bt1G327	562-109	g2340013	391	417	4.90E-39	55	98	(Z98682) YlbP protein [Bacillus subtilis]
271	Bt1Gc465	Bt1G328	648-825	g1770012	203	209	1.30E-16	70	16	(Z75208) hypothetical protein [Bacillus subtilis]
272	Bt1Gc470	Bt1G329	845-1	g2828499	1058	1004	3.10E-101	71	75	HYPOTHETICAL 41.8 KD PROTEIN (ORFM) []
273	Bt1Gc471	Bt1G330	742-79	g2619050	316	325	2.80E-29	46	61	(AF027868) YocH [Bacillus subtilis]
274	Bt1Gc475	Bt1G331	493-1	g4584148	811	813	5.40E-81	94	17	(AJ010111) pyruvate carboxylase [Bacillus cereus]
275	Bt1Gc476	Bt1G332	175-1	g2108269	116	133	6.10E-09	41	41	(X92868) mercuric resistance operon regulatory protein [Bacillus subtilis]
275	Bt1Gc476	Bt1G333	918-462	g4115379	105	160	2.00E-11	30	59	(AC005967) putative carbonyl reductase [Arabidopsis thaliana]
276	Bt1Gc467	Bt1G334	599-162	g2506860	245	188	9.10E-15	40	66	AQUAPORIN Z (BACTERIAL NODULIN-LIKE INTRINSIC PROTEIN) [Escherichia coli]
277	Bt1Gc477	Bt1G335	1-464	g3915396	688	693	2.80E-68	82	52	SUCCINYL-COA SYNTHETASE ALPHA CHAIN (SCS-ALPHA) (VEGETATIVE PROTEIN 239) (VEG239) [Bacillus subtilis]
278	Bt1Gc472	Bt1G336	1-628	g3123227	913	938	3.00E-94	84	41	GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
279	Bt1Gc479	Bt1G337	1-371	g1881262	345	330	8.20E-30	59	32	SYNTHETASE [Bacillus subtilis]
280	Bt1Gc481	Bt1G338	903-598	g1788774	82	137	2.30E-09	28	54	(AB001488) SIMILAR TO ACYL-COA DEHYDROGENASE. [Bacillus subtilis]
281	Bt1Gc482	Bt1G339	841-1	g2127339	493	517	1.20E-48	36	27	(AE000330) orf, hypothetical protein [Escherichia coli]
282	Bt1Gc484	Bt1G340	608-1036	g2462122	345	362	3.30E-33	54	99	collagenase - Clostridium histolyticum [Clostridium histolyticum]
283	Bt1Gc483	Bt1G341	857-598	g2226142	164	134	4.80E-09	52	52	(Y11141) microbial collagenase precursor [Bacillus cereus]
284	Bt1Gc485	Bt1G342	1-230	g135723	162	129	7.50E-08	49	23	hypothetical protein [Bacillus subtilis]
284	Bt1Gc485	Bt1G343	480-836	g2274943	281	286	3.80E-25	45	79	THREONINE DEHYDRATASE CATABOLIC (THREONINE DEAMINASE) [Escherichia coli]
285	Bt1Gc486	Bt1G344	1-240	g2635731	340	344	2.70E-31	75	23	(AJ000346) NapB protein [Enterococcus hirae]
285	Bt1Gc486	Bt1G345	749-280	g2635732	218	291	1.10E-25	36	61	alternate gene name: yutA [Bacillus subtilis]
286	Bt1Gc478	Bt1G346	1001-293	g2632020	462	553	1.90E-53	43	76	(Z99120) yunB [Bacillus subtilis]
286	Bt1Gc478	Bt1G347	1001-800	g4514328	171	187	1.20E-14	48	40	(AJ002571) Ykfc [Bacillus subtilis]
287	Bt1Gc488	Bt1G348	100-468	g2633028	341	356	1.40E-32	55	99	(AB013368) Ykfc [Bacillus halodurans]
288	Bt1Gc489	Bt1G349	865-129	g2635751	186	230	3.20E-19	62	22	(Z99107) yetG [Bacillus subtilis]
289	Bt1Gc490	Bt1G350	166-1	g2635737	142	161	6.60E-12	55	48	(Z99120) similar to ribonuclease [Bacillus subtilis]
290	Bt1Gc495	Bt1G351	288-1	g1881243	112	147	2.00E-10	32	54	(Z99120) yunG [Bacillus subtilis]
291	Bt1Gc497	Bt1G352	395-1109	g2633719	357	404	1.20E-37	40	97	(AB001488) FUNCTION UNKNOWN, CONTAINS MUTT MOTIF, SIMILAR PRODUCTS IN SYNECHOCYSTIS. [Bacillus subtilis]
292	Bt1Gc498	Bt1G353	709-232	g3320120	237	271	1.50E-23	35	41	(Z99111) ykrK [Bacillus subtilis]
293	Bt1Gc496	Bt1G354	1-551	g1731364	407	401	2.40E-37	50	46	(U66669) 3-hydroxyisobutyryl-coenzyme A hydrolase [Homo sapiens]
294	Bt1Gc499	Bt1G355	1-779	g585372	330	267	1.60E-22	41	39	HYPOTHETICAL PROTEASE IN ROCR-PURA INTERGENIC REGION [Bacillus subtilis]
295	Bt1Gc500	Bt1G356	698-41	g671632	128	210	4.20E-17	26	99	PYRUVATE KINASE [Lactococcus lactis]
296	Bt1Gc502	Bt1G357	911-293	g4098413	95	173	1.20E-12	32	32	(Z16422) unknown [Staphylococcus aureus]
297	Bt1Gc505	Bt1G358	402-832	g606198	308	260	8.50E-22	49	29	(U77495) putative integrase [Leuconostoc oenos bacteriophage 10MC]
298	Bt1Gc494	Bt1G359	571-1	g1684651	383	298	2.00E-26	38	48	(U18997) pantothenate permease [Escherichia coli]
298	Bt1Gc494	Bt1G360	886-593	g1684652	163	128	2.10E-08	34	71	(Z82987) unknown similar to quinolon resistance protein NorA [Bacillus subtilis]
										(Z82987) unknown, similar to cytolysin SlyA from

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
299	Bt1Gc509	Bt1G361	Jan-91	g4584128	120	136	2.90E-09	94	45	Salmonella typhimurium [Bacillus subtilis]
299	Bt1Gc509	Bt1G362	198-814	g3123214	801	694	2.20E-68	75	70	(AJ010135) ribosomal protein S2 [Bacillus cereus]
300	Bt1Gc511	Bt1G363	182-496	g1176769	448	454	5.90E-43	80	99	ELONGATION FACTOR TS (EF-TS) [Bacillus subtilis]
300	Bt1Gc511	Bt1G364	560-994	g1176770	554	549	5.10E-53	71	99	HYPOTHETICAL 12.3 KD PROTEIN IN CWLA-CISA INTERGENIC REGION [Bacillus subtilis]
300	Bt1Gc511	Bt1G365	1016-1205	g1176771	266	212	4.50E-17	83	18	HYPOTHETICAL 16.6 KD PROTEIN IN CWLA-CISA INTERGENIC REGION [Bacillus subtilis]
301	Bt1Gc512	Bt1G366	1-1052	g1724012	145	227	1.80E-17	32	14	HYPOTHETICAL 38.3 KD PROTEIN IN CWLA-CISA INTERGENIC REGION [Bacillus subtilis]
301	Bt1Gc512	Bt1G367	524-815	g121841	77	139	1.10E-08	28	23	HYPOTHETICAL 132.7 KD PROTEIN IN CSPB-GLPP INTERGENIC REGION [Bacillus subtilis]
302	Bt1Gc513	Bt1G368	1-1082	g548931	1161	579	1.50E-93	76	92	ENDOGLUCANASE PRECURSOR (ENDO-1,4-BETA-GLUCANASE) (CELLULASE) [Clostridium acetobutylicum]
303	Bt1Gc514	Bt1G369	202-1	g141422	106	119	1.90E-07	38	80	S-LAYER PROTEIN PRECURSOR (SURFACE LAYER PROTEIN) [Bacillus thuringiensis]
304	Bt1Gc504	Bt1G370	1-926	g416901	981	973	5.90E-98	62	89	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN SMAI RESTRICTION SYSTEM 5'REGION [Serratia marcescens]
305	Bt1Gc515	Bt1G371	1-839	g2635858	339	442	5.80E-41	34	36	ASPARTATE-SEMIALDEHYDE DEHYDROGENASE (ASA DEHYDROGENASE) [Bacillus subtilis]
306	Bt1Gc503	Bt1G372	497-1	g2266425	638	645	3.40E-63	72	37	(Z99121) yvgS [Bacillus subtilis]
307	Bt1Gc517	Bt1G373	648-101	g1730917	507	536	1.20E-51	53	75	(Y13917) yngH [Bacillus subtilis]
308	Bt1Gc521	Bt1G375	423-1	g2833192	307	328	1.30E-29	40	80	HYPOTHETICAL 27.7 KD PROTEIN IN GPSA-SPOIVA INTERGENIC REGION (ORF1) [Bacillus subtilis]
309	Bt1Gc523	Bt1G376	503-87	g118721	406	258	2.70E-21	74	21	DINB PROTEIN [Bacillus subtilis]
310	Bt1Gc526	Bt1G377	707-225	g112703	619	533	2.50E-51	74	99	DNAB PROTEIN (HEAT SHOCK PROTEIN 70) (HSP70) [Bacillus megaterium]
311	Bt1Gc528	Bt1G378	1-735	g538771	779	795	4.30E-79	59	51	HYPOTHETICAL 17.8 KD PROTEIN IN SERS-DNAH INTERGENIC REGION [Bacillus subtilis]
312	Bt1Gc529	Bt1G379	801-379	g4753857	121	156	6.70E-11	27	45	lysine-specific permease - Escherichia coli [] (AL049754) putative ABC transporter ATP-binding protein [Streptomyces coelicolor]
313	Bt1Gc527	Bt1G380	1089-1	g1124822	1064	1074	1.20E-108	54	76	unknown protein [Bacillus thuringiensis]
314	Bt1Gc533	Bt1G382	1-975	g2811132	1046	1011	5.60E-102	61	60	(AF013216) Mls [Myxococcus xanthus]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
314	Bt1Gc533	Bt1G383	769-975	g167521	118	78	0.0041	39	74	(M16219) glyoxysomal malate synthase (EC 4.1.3.1) [Cucumis sativus]
315	Bt1Gc537	Bt1G384	1-385	g113333	377	383	2.00E-35	52	61	METHYLPHOSPHOTRIESTER-DNA ALKYLTRANSFERASE [Bacillus subtilis]
315	Bt1Gc537	Bt1G385	381-880	g113335	337	363	2.60E-33	44	93	METHYLATED-DNA--PROTEIN-CYSTEINE METHYLTRANSFERASE (O-6-METHYLGUANINE-DNA ALKYLTRANSFERASE) [Bacillus subtilis]
316	Bt1Gc538	Bt1G386	881-1	g2507017	990	1027	1.10E-103	68	78	HEMOLYSIN BL BINDING COMPONENT PRECURSOR (ENTEROTOXIN 40 KD SUBUNIT) [Bacillus cereus]
317	Bt1Gc540	Bt1G387	1-501	g1172036	533	485	3.10E-46	63	38	XANTHINE PERMEASE [Bacillus subtilis]
317	Bt1Gc540	Bt1G388	676-532	g418336	95	112	6.00E-06	32	14	HYPOTHETICAL 39.9 KD PROTEIN IN AMYLASE 3'REGION [Bacillus acidopullulicus]
318	Bt1Gc522	Bt1G389	489-187	g2635865	351	356	1.40E-32	63	99	(Z99121) similar to hypothetical proteins [Bacillus subtilis]
319	Bt1Gc542	Bt1G390	682-1	g118333	554	566	8.00E-55	50	41	INDOLE-3-PYRUVATE DECARBOXYLASE (INDOLEPYRUVATE DECARBOXYLASE) [Enterobacter cloacae]
320	Bt1Gc539	Bt1G391	1-100	g2984723	125	141	8.70E-10	79	46	(AF053927) probable spore germination protein F [Bacillus cereus]
320	Bt1Gc539	Bt1G392	451-1019	g2501575	282	280	1.60E-24	33	62	HYPOTHETICAL 33.3 KD PROTEIN SLL1263 [Synechocystis sp.]
321	Bt1Gc543	Bt1G393	423-1	g1750129	282	311	8.40E-28	40	92	(U66480) YncE [Bacillus subtilis]
321	Bt1Gc543	Bt1G394	845-640	g2226190	182	202	3.00E-16	51	47	(Y14081) hypothetical protein [Bacillus subtilis]
322	Bt1Gc544	Bt1G395	1-711	g1881236	338	443	3.30E-41	42	31	(AB001488) PROBABLE DNA TOPOISOMERASE III [Bacillus subtilis]
323	Bt1Gc546	Bt1G396	1-847	g3822300	115	274	8.60E-23	23	36	(AF081810) viral enhancing factor 1 [Lymantria dispar nucleopolyhedrovirus]
324	Bt1Gc547	Bt1G397	1-279	g2293285	272	248	4.00E-21	56	60	(AF008220) YtaB [Bacillus subtilis]
325	Bt1Gc548	Bt1G398	876-185	g2619029	250	294	5.30E-26	31	76	(AF027868) YoaR [Bacillus subtilis]
326	Bt1Gc549	Bt1G399	1-310	g1434920	184	238	6.90E-19	43	13	(D86346) crystal protein [Bacillus thuringiensis]
327	Bt1Gc551	Bt1G400	1-440	g3308998	451	464	5.20E-44	60	75	(AB015998) CBP21 precursor [Serratia marcescens]
327	Bt1Gc551	Bt1G401	661-784	g586817	165	181	9.50E-14	76	13	HYPOTHETICAL 34.0 KD PROTEIN IN COTF-TETB INTERGENIC REGION [Bacillus subtilis]
328	Bt1Gc550	Bt1G402	615-1	g728913	339	372	3.40E-33	43	21	MG(2+) TRANSPORT ATPASE, P-TYPE 1 [Escherichia coli]
329	Bt1Gc553	Bt1G403	901-523	g2497392	638	640	1.20E-62	97	50	INSERTION SEQUENCE IS232 PUTATIVE ATP-

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
330	Bt1Gc552	Bt1G404	391-612	g225559	368	368	7.70E-34	99	15	BINDING PROTEIN [Insertion sequence IS232]
331	Bt1Gc554	Bt1G405	833-165	g2650107	140	232	2.00E-19	24	76	ORF IS231C [Bacillus thuringiensis] conserved hypothetical protein [Archaeoglobus fulgidus]
332	Bt1Gc556	Bt1G406	1-889	g3868868	320	372	2.90E-34	30	63	(AB014075) Orf5u [Clostridium histolyticum]
333	Bt1Gc555	Bt1G407	1072-91	g1321706	564	699	6.50E-69	44	78	(X95309) triacylglycerol lipase [Bacillus thermocatenulatus]
334	Bt1Gc557	Bt1G408	1148-233	g1730911	924	890	3.70E-89	62	82	HYPOTHETICAL 42.6 KD PROTEIN IN BSAA-IL VD INTERGENIC REGION [Bacillus subtilis]
335	Bt1Gc560	Bt1G409	920-115	g1945088	746	760	2.20E-75	51	70	(D88802) S. lividans glycosyl transferase; J50636 (311) [Bacillus subtilis]
336	Bt1Gc559	Bt1G410	503-1	g473953	83	168	1.20E-11	27	30	(D29673) hydrolase [Lactobacillus sp.]
337	Bt1Gc563	Bt1G411	775-515	g141373	173	177	1.30E-13	44	99	SIGMA-K FACTOR PROCESSING REGULATORY PROTEIN BOFA (BYPASS-OF-FORESPORE PROTEIN) [Bacillus subtilis]
338	Bt1Gc565	Bt1G412	665-1096	g3024926	178	222	2.30E-18	35	61	HYPOTHETICAL PROTEIN MJ0912 [Methanococcus jannaschii]
339	Bt1Gc567	Bt1G413	600-869	g3451437	74	128	2.40E-07	28	17	(AL031350) putative secreted protein [Streptomyces coelicolor]
340	Bt1Gc562	Bt1G414	1-1074	g3183185	1418	1349	8.50E-138	80	58	GTP-BINDING PROTEIN TYPA/BIPA HOMOLOG [Bacillus subtilis]
341	Bt1Gc568	Bt1G415	1-395	g3183519	340	375	1.40E-34	58	30	GID PROTEIN [Bacillus subtilis]
341	Bt1Gc568	Bt1G416	458-1026	g2126912	534	471	9.30E-45	52	63	codV protein - Bacillus subtilis []
342	Bt1Gc570	Bt1G417	1-136	g2634032	85	117	1.90E-06	55	13	(Z99112) nusaA [Bacillus subtilis]
342	Bt1Gc570	Bt1G418	148-420	g418461	290	252	1.50E-21	59	99	HYPOTHETICAL 10.4 KD PROTEIN IN NUSA-INFB INTERGENIC REGION (ORF3) [Bacillus subtilis]
342	Bt1Gc570	Bt1G419	418-717	g418462	309	318	1.50E-28	61	99	PROBABLE RIBOSOMAL PROTEIN IN NUSA-INFB INTERGENIC REGION (ORF4) [Bacillus subtilis]
343	Bt1Gc541	Bt1G420	621-374	g266724	65	141	7.50E-09	37	18	PROTEIN P60 PRECURSOR (INVASION-ASSOCIATED PROTEIN) [Listeria innocua]
344	Bt1Gc569	Bt1G421	266-44	g2633659	114	132	7.80E-09	35	99	(Z99110) ykzH [Bacillus subtilis]
345	Bt1Gc572	Bt1G422	161-1	g1881335	97	111	5.50E-06	44	19	(AB001488) SIMILAR TO YQFU, YXKD, YITB OF B. SUBTILIS. [Bacillus subtilis]
345	Bt1Gc572	Bt1G423	1443-1338	g80073	137	153	1.40E-10	92	11	beta-lactamase (EC 3.5.2.6) precursor - Bacillus cereus [Bacillus cereus]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvr	NCBI gi description
346	Bt1Gc574	Bt1G424	1386-101	g2094843	711	499	1.00E-47	39	89	(Z95324) mgfE [Mycobacterium tuberculosis]
347	Bt1Gc573	Bt1G425	910-1	g1354814	229	356	1.40E-32	29	79	(U57065) ZmaR [Bacillus cereus]
348	Bt1Gc571	Bt1G426	1-762	g1346893	226	154	2.40E-14	28	28	PTS SYSTEM, GLUCOSE-SPECIFIC IIABC COMPONENT (EIIABC-GLC) (GLUCOSE-PERMEASE IIABC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, ABC COMPONENT) (EII-GLC / EIII-GLC) [Mycoplasma genitalium]
349	Bt1Gc577	Bt1G427	1-1424	g1731087	185	233	1.60E-19	37	41	HYPOTHETICAL 39.0 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
350	Bt1Gc575	Bt1G428	1016-1	g3915990	395	375	1.40E-34	36	46	HYPOTHETICAL 58.3 KD PROTEIN IN GLPD-CSPB INTERGENIC REGION [Bacillus subtilis]
351	Bt1Gc578	Bt1G429	361-1	g3123299	572	480	1.00E-45	88	26	HYPOTHETICAL TRANSPORT PROTEIN IN NDHF-CSGA INTERGENIC REGION (ORF1) [Bacillus subtilis]
352	Bt1Gc566	Bt1G430	538-38	g3915543	566	566	8.00E-55	62	99	HYPOTHETICAL 18.9 KD PROTEIN IN CYP-AADK INTERGENIC REGION [Bacillus subtilis]
353	Bt1Gc579	Bt1G431	1-1009	g3322837	196	375	4.70E-34	27	58	(AE001229) T. pallidum predicted coding region TP0544 [Treponema pallidum]
354	Bt1Gc583	Bt1G432	839-1	g2507445	881	877	8.90E-88	61	63	D-SERINE DEHYDRATASE (D-SERINE DEAMINASE) [Escherichia coli]
355	Bt1Gc584	Bt1G433	252-1031	g2661703	138	231	2.50E-19	27	82	(AL009204) hypothetical protein SC9B10.17 [Streptomyces coelicolor]
356	Bt1Gc585	Bt1G434	1072-132	g732405	105	190	4.50E-14	31	24	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN FEUA-SIGW INTERGENIC REGION (ORF3) [Bacillus subtilis]
357	Bt1Gc582	Bt1G435	1-165	g1575578	114	148	1.00E-09	53	13	(U67196) histidine protein kinase [Thermotoga maritima]
357	Bt1Gc582	Bt1G436	278-1101	g2634226	862	875	1.40E-87	59	83	(Z99113) similar to alcohol dehydrogenase [Bacillus subtilis]
358	Bt1Gc590	Bt1G437	1-275	g2634266	267	276	4.30E-24	62	57	(Z99114) yoaS [Bacillus subtilis]
358	Bt1Gc590	Bt1G438	289-540	g2634267	308	322	5.70E-29	76	99	(Z99114) similar to transcriptional regulator [Bacillus subtilis]
359	Bt1Gc588	Bt1G439	1-798	g121467	1081	420	1.60E-71	81	63	PROTON/SODIUM-GLUTAMATE SYMPORT PROTEIN (GLUTAMATE-ASPARTATE CARRIER PROTEIN) [Bacillus stearothermophilus]
360	Bt1Gc591	Bt1G440	480-1	g2909586	98	137	2.30E-09	26	76	(AL021927) hypothetical protein Rv0158 [Mycobacterium tuberculosis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
361	Bt1Gc593	Bt1G441	1030-35	g3861147	199	215	4.30E-17	27	60	(AJ235272) BICYCLOMYCIN RESISTANCE PROTEIN (bcr1) [Rickettsia prowazekii]
362	Bt1Gc592	Bt1G442	1182-336	g2621332	187	234	1.20E-19	34	99	(AE000813) conserved protein [Methanobacterium thermoautotrophicum]
363	Bt1Gc596	Bt1G443	1-431	g2577968	281	241	2.20E-20	37	84	(Y15254) SapB protein [Bacillus subtilis]
363	Bt1Gc596	Bt1G444	534-715	g2635584	80	108	2.70E-06	38	35	(Z99119) yuaI [Bacillus subtilis]
364	Bt1Gc597	Bt1G445	839-20	g586703	585	510	6.90E-49	41	68	HYPOTHETICAL 43.0 KD PROTEIN IN PROK-TAG INTERGENIC REGION (F402) [Escherichia coli]
365	Bt1Gc599	Bt1G446	305-724	g2633752	177	248	4.00E-21	38	99	(Z99111) ykvS [Bacillus subtilis]
366	Bt1Gc601	Bt1G447	1-618	g548909	519	532	3.20E-51	52	48	SERINE TRANSPORTER [Escherichia coli]
367	Bt1Gc600	Bt1G448	807-1	g3914084	518	545	1.90E-52	40	34	MUTS2 PROTEIN [Bacillus subtilis]
368	Bt1Gc606	Bt1G450	784-80	g1168824	864	588	3.70E-57	65	100	CYTOCHROME C-TYPE BIOGENESIS PROTEIN CCDA [Bacillus subtilis]
369	Bt1Gc603	Bt1G451	1058-1	g628966	509	544	1.70E-52	40	49	hypothetical protein gamma - Streptococcus pyogenes plasmid pDB101 [Streptococcus pyogenes]
370	Bt1Gc607	Bt1G452	247-1	g3322943	215	247	2.90E-20	51	16	(AE001239) lysyl-tRNA synthetase (lysS-1) [Treponema pallidum]
371	Bt1Gc608	Bt1G453	853-174	g2619012	79	119	1.10E-06	49	14	(AF027868) fatty acid desaturase [Bacillus subtilis]
372	Bt1Gc611	Bt1G454	1545-479	g2116973	1253	1289	1.90E-131	67	100	(D87979) YfkK [Bacillus subtilis]
372	Bt1Gc611	Bt1G455	1996-1665	g1731305	174	131	1.00E-08	45	88	HYPOTHETICAL 14.7 KD PROTEIN IN IDH-DEOR INTERGENIC REGION [Bacillus subtilis]
373	Bt1Gc605	Bt1G456	1-567	g4155188	586	532	3.20E-51	58	41	(AE001495) short-chain fatty acids transporter [Helicobacter pylori J99]
374	Bt1Gc613	Bt1G457	416-1	g400802	521	520	6.00E-50	68	55	PHOSPHOGLYCERATE MUTASE 1 (PHOSPHOGLYCEROMUTASE 1) (PGAM 1) (BPG-DEPENDENT PGAM 1) [Escherichia coli]
374	Bt1Gc613	Bt1G458	1037-588	g400755	248	260	2.10E-22	36	89	PETP PROTEIN [Rhodobacter capsulatus]
375	Bt1Gc612	Bt1G459	1-443	g586898	312	292	8.70E-26	41	80	HYPOTHETICAL 21.0 KD PROTEIN IN LYSS-MECB INTERGENIC REGION [Bacillus subtilis]
376	Bt1Gc614	Bt1G460	1-1251	g1175719	758	799	1.60E-79	40	64	PENICILLIN-BINDING PROTEIN 3 (PBP 3) [Bacillus subtilis]
377	Bt1Gc617	Bt1G461	486-887	g4154073	240	265	6.30E-23	42	99	(AL035161) conserved hypothetical protein [Streptomyces coelicolor]
378	Bt1Gc616	Bt1G462	1083-1	g1171759	249	452	9.60E-43	31	62	BACILLOLYSIN PRECURSOR (NEUTRAL PROTEASE) [Bacillus brevis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
378	Bt1Gc616	Bt1G463	1083-1	g2239276	264	399	4.00E-37	32	63	(L77763) neutral protease A [Bacillus thuringiensis]
379	Bt1Gc619	Bt1G464	919-484	g2497400	184	229	4.10E-19	38	60	HYPOTHETICAL TRANSPOSASE-LIKE PROTEIN HI1721 [Haemophilus influenzae Rd]
380	Bt1Gc623	Bt1G465	74-926	g1723610	441	530	5.20E-51	39	94	HYPOTHETICAL 31.5 KD PROTEIN IN GLVBC 3'REGION [Bacillus subtilis]
381	Bt1Gc622	Bt1G466	899-422	g1653921	227	232	2.00E-19	33	82	(D90917) hypothetical protein [Synecocystis sp.]
382	Bt1Gc624	Bt1G467	529-651	g1177016	118	140	6.90E-09	61	10	HYPOTHETICAL 43.7 KD PROTEIN IN KATB 3'REGION [Bacillus subtilis]
383	Bt1Gc625	Bt1G468	778-402	g2497400	171	211	3.30E-17	36	57	HYPOTHETICAL TRANSPOSASE-LIKE PROTEIN HI1721 [Haemophilus influenzae Rd]
384	Bt1Gc627	Bt1G469	1-223	g1176771	306	257	4.40E-22	85	21	HYPOTHETICAL 38.3 KD PROTEIN IN CWLA-CISA INTERGENIC REGION [Bacillus subtilis]
384	Bt1Gc627	Bt1G470	252-563	g1168520	409	419	3.00E-39	77	73	PUTATIVE ARSENATE REDUCTASE [Bacillus subtilis]
385	Bt1Gc620	Bt1G471	1-926	g4455080	176	168	2.00E-12	29	63	(AF119621) Dtl [Pseudomonas abietaniphila]
385	Bt1Gc620	Bt1G472	926-559	g3688811	325	348	1.00E-31	53	33	(AF084104) maltose transporter ATP-binding protein [Bacillus firmus]
386	Bt1Gc628	Bt1G473	1-434	g586865	134	136	1.80E-08	25	37	HYPOTHETICAL 43.8 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION [Bacillus subtilis]
387	Bt1Gc629	Bt1G474	1-182	g2896709	132	121	3.90E-07	46	22	(AL021897) hypothetical protein Rv1072 [Mycobacterium tuberculosis]
387	Bt1Gc629	Bt1G475	1199-904	g549113	337	296	7.50E-26	72	21	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS231E [Bacillus thuringiensis]
388	Bt1Gc631	Bt1G476	1-433	g3025311	256	237	5.80E-20	35	74	HYPOTHETICAL 21.7 KD PROTEIN IN NRGB-SPOIIQ INTERGENIC REGION [Bacillus subtilis]
389	Bt1Gc630	Bt1G477	1-1038	g4127525	1277	1205	1.50E-122	70	88	(AJ011526) D-stereospecific peptide hydrolase [Bacillus cereus]
390	Bt1Gc610	Bt1G478	520-1285	g1945654	375	424	8.90E-40	35	100	(Z94043) hypothetical protein [Bacillus subtilis]
391	Bt1Gc634	Bt1G479	1049-686	g225559	129	215	7.00E-17	50	21	ORF IS231C [Bacillus thuringiensis]
392	Bt1Gc636	Bt1G480	195-845	g2635861	395	449	2.00E-42	38	100	(Z99121) similar to hypothetical proteins [Bacillus subtilis]
393	Bt1Gc632	Bt1G481	223-382	g1665854	85	123	7.00E-08	41	37	(Z82015) yukJ [Bacillus subtilis]
394	Bt1Gc637	Bt1G482	867-739	g225559	195	195	1.10E-14	91	9	ORF IS231C [Bacillus thuringiensis]
395	Bt1Gc635	Bt1G483	1-1010	g2226258	616	619	1.90E-60	42	78	(Y14084) hypothetical protein [Bacillus subtilis]
396	Bt1Gc640	Bt1G484	288-756	g1945109	78	147	2.00E-10	26	99	(D88802) H. influenzae, ribosomal protein alanine acetyltransferase; P44305 (189) [Bacillus subtilis]
396	Bt1Gc640	Bt1G485	1-1141	g2415396	120	121	3.70E-07	40	34	(AF015775) carboxypeptidase [Bacillus subtilis]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
397	Bt1Gc642	Bt1G486	953-286	g1705428	604	605	5.90E-59	59	84	BACITRACIN RESISTANCE PROTEIN (PUTATIVE UNDECAPRENOL KINASE) [Escherichia coli]
398	Bt1Gc644	Bt1G487	1-396	g2127182	141	202	9.10E-16	29	37	rap60 protein - Bacillus subtilis plasmid pTA1060 [Bacillus subtilis]
398	Bt1Gc644	Bt1G488	968-625	g2497392	570	584	9.90E-57	98	46	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
399	Bt1Gc645	Bt1G489	1-702	g1176282	621	653	4.80E-64	54	67	HYPOTHETICAL OXIDOREDUCTASE IN GNTR-GGT INTERGENIC REGION [Escherichia coli]
400	Bt1Gc646	Bt1G490	616-1	g2226164	435	461	1.10E-43	44	99	(Y14080) hypothetical protein [Bacillus subtilis]
401	Bt1Gc648	Bt1G491	Jan-66	g3122932	111	111	9.10E-06	86	6	QUEUEINE TRNA-RIBOSYLTRANSFERASE (TRNA-GUANINE TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME) [Bacillus subtilis]
401	Bt1Gc648	Bt1G492	94-357	g2635235	265	196	1.30E-15	62	99	(Z99118) similar to hypothetical proteins [Bacillus subtilis]
401	Bt1Gc648	Bt1G493	766-391	g2635234	228	238	4.60E-20	33	83	(Z99118) yrzE [Bacillus subtilis]
402	Bt1Gc643	Bt1G494	442-1	g3599371	402	348	1.00E-31	57	65	(AF082668) CsrR [Streptococcus pyogenes]
403	Bt1Gc650	Bt1G495	21-1317	g1723608	200	140	1.10E-09	45	78	HYPOTHETICAL PROTEIN IN GLVBC 3 REGION [Bacillus subtilis]
404	Bt1Gc649	Bt1G496	948-109	g2624002	679	664	3.30E-65	49	100	(Z93941) YuxA [Bacillus subtilis]
405	Bt1Gc647	Bt1G497	1138-919	g2688589	173	199	6.20E-16	49	28	(AE001167) phosphoglycerate mutase (gpmA) [Borrelia burgdorferi]
406	Bt1Gc655	Bt1G498	262-1	g3256746	113	135	3.80E-09	43	76	(AP000002) 109aa long hypothetical protein [Pyrococcus horikoshii]
407	Bt1Gc654	Bt1G499	1-473	g1345692	509	514	2.60E-49	53	72	CHLORAMPHENICOL ACETYLTRANSFERASE (CAT) [Clostridium butyricum]
408	Bt1Gc656	Bt1G500	678-346	g586039	313	349	7.90E-32	59	64	SINGLE-STRAND BINDING PROTEIN (SSB) (HELIX-DESTABILIZING PROTEIN) [Bacillus subtilis]
409	Bt1Gc657	Bt1G501	721-1	g231698	1086	1091	1.90E-110	80	53	CATALASE [ ]
410	Bt1Gc658	Bt1G502	915-676	g2127290	57	123	1.80E-07	34	32	transposase (insertion sequence IS231) - Bacillus thuringiensis [Bacillus thuringiensis]
410	Bt1Gc658	Bt1G503	1336-1194	g2507017	126	156	1.10E-10	60	13	HEMOLYSIN BL BINDING COMPONENT PRECURSOR (ENTEROTOXIN 40 KD SUBUNIT) [Bacillus cereus]
411	Bt1Gc652	Bt1G504	1-417	g1346454	65	130	5.40E-08	31	28	LIPASE 1 PRECURSOR (TRIACYLGLYCEROL LIPASE) [Psychrobacter immobilis]
412	Bt1Gc653	Bt1G505	335-1	g3915460	357	364	2.00E-33	58	51	PUTATIVE NAD(P)H NITROREDUCTASE YFKO [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
412	Bt1Gc653	Bt1G506	485-838	g586816	373	398	5.10E-37	61	92	HYPOTHETICAL 14.7 KD PROTEIN IN COTF-TETB INTERGENIC REGION [Bacillus subtilis]
413	Bt1Gc661	Bt1G507	462-1	g1730956	540	540	4.60E-52	64	85	HYPOTHETICAL 21.1 KD PROTEIN IN COTD-KDUD INTERGENIC REGION [Bacillus subtilis]
413	Bt1Gc661	Bt1G508	948-684	g116956	137	189	7.10E-15	47	99	SPORE COAT PROTEIN D [Bacillus subtilis]
414	Bt1Gc662	Bt1G509	1-201	g2226145	149	171	4.40E-12	50	13	(Y14079) hypothetical protein [Bacillus subtilis]
414	Bt1Gc662	Bt1G510	1134-765	g1263187	102	161	1.40E-11	30	43	(U24215) HOMODA hydrolase [Pseudomonas putida]
415	Bt1Gc664	Bt1G511	1-1090	g2415704	81	123	5.20E-07	32	16	(AJ001361) proline iminopeptidase [Propionibacterium freudenreichii subsp. shermanii]
416	Bt1Gc665	Bt1G512	1-759	g2497382	537	570	4.40E-57	85	29	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
416	Bt1Gc665	Bt1G513	759-455	g223559	490	501	6.20E-48	98	21	ORF IS231C [Bacillus thuringiensis]
417	Bt1Gc666	Bt1G514	11-927	g2493595	512	558	5.60E-54	38	92	PUTATIVE POTASSIUM CHANNEL PROTEIN MJ0138.1 [Methanococcus jannaschii]
418	Bt1Gc670	Bt1G515	393-16	g1763703	193	229	4.10E-19	37	99	(Z83337) similar to hydroxymyristoyl-(acyl carrier protein) dehydratase [Bacillus subtilis]
418	Bt1Gc670	Bt1G516	1586-469	g1001578	752	784	6.30E-78	39	95	(D64000) aspartate aminotransferase [Synechocystis sp.]
419	Bt1Gc671	Bt1G517	1-548	g2340009	463	481	8.10E-46	48	44	(Z98682) YlbM protein [Bacillus subtilis]
420	Bt1Gc672	Bt1G518	727-1014	g465650	192	176	1.70E-13	38	99	HYPOTHETICAL 11.0 KD PROTEIN IN HSP18 3'REGION (ORF1) [Clostridium acetobutylicum]
421	Bt1Gc673	Bt1G519	1060-1	g1174634	93	168	2.10E-12	26	60	TELLURITE RESISTANCE PROTEIN TEHB HOMOLOG [Haemophilus influenzae Rd]
422	Bt1Gc674	Bt1G520	26-280	g418608	212	223	1.80E-18	48	99	HYPOTHETICAL 9.8 KD PROTEIN IN SPOVFA 5'REGION (ORFZ) [Bacillus subtilis]
422	Bt1Gc674	Bt1G521	448-1218	g417830	814	814	4.20E-81	60	86	DIPICOLINATE SYNTHASE, A CHAIN [Bacillus subtilis]
423	Bt1Gc675	Bt1G522	1-1038	g2633162	329	245	7.60E-26	25	83	(Z99108) similar to multidrug resistance protein [Bacillus subtilis]
424	Bt1Gc676	Bt1G523	1-756	g585648	500	546	1.10E-52	39	35	PENICILLIN-BINDING PROTEIN 2B (PBP-2B) [Bacillus subtilis]
424	Bt1Gc676	Bt1G524	1-752	g2149903	241	348	8.20E-31	33	33	(U94707) penicillin-binding protein [Enterococcus faecalis]
425	Bt1Gc679	Bt1G525	965-621	g3116222	140	124	3.90E-07	28	29	(AB007122) transporter [Arthrobacter sp.]
426	Bt1Gc681	Bt1G526	1-341	g1945711	277	291	1.10E-25	49	27	(Z94043) hypothetical protein [Bacillus subtilis]
426	Bt1Gc681	Bt1G527	463-748	g2635928	274	308	1.70E-27	59	23	(Z99121) similar to maltodextrin transport system permease [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
427	Bt1Gc682	Bt1G528	1365-620	g1620923	629	572	1.90E-55	49	88	(Z79580) putative ORF [Bacillus subtilis]
428	Bt1Gc680	Bt1G529	564-1	g2612894	135	212	8.10E-17	25	47	(AF015825) hypothetical glycosyl transferase [Bacillus subtilis]
429	Bt1Gc683	Bt1G530	301-1	g1945088	261	273	9.00E-24	51	26	(D88802) S. lividans glycosyl transferase; JS0636 (311) [Bacillus subtilis]
429	Bt1Gc683	Bt1G531	1042-606	g4033455	277	254	9.20E-22	41	81	SIGNAL PEPTIDASE I P (SPASE I) (LEADER PEPTIDASE I) [Bacillus subtilis]
430	Bt1Gc684	Bt1G532	1-1072	g2293165	1264	1269	2.60E-129	66	56	(AF008220) asparagine synthase [Bacillus subtilis]
431	Bt1Gc685	Bt1G533	588-1	g3907604	271	299	1.60E-26	36	77	(AF094575) putative N-acetyl-mannosamine transferase Cps19aF [Streptococcus pneumoniae]
431	Bt1Gc685	Bt1G534	1554-619	g1730929	187	248	4.80E-21	29	83	HYPOTHETICAL 42.0 KD PROTEIN IN DAPB-PAPS INTERGENIC REGION [Bacillus subtilis]
432	Bt1Gc677	Bt1G535	15-278	g131533	316	287	2.90E-25	72	99	PHOSPHOCARRIER PROTEIN HPR (HISTIDINE-CONTAINING PROTEIN) [Bacillus subtilis]
432	Bt1Gc677	Bt1G536	281-780	g3021327	641	480	1.00E-45	77	29	(AJ005075) enzyme 1 [Bacillus megaterium]
433	Bt1Gc686	Bt1G537	281-1	g3256836	76	136	1.10E-08	36	30	(AP000002) 304aa long hypothetical protein [Pyrococcus horikoshii]
434	Bt1Gc687	Bt1G538	1199-628	g2648503	175	272	1.10E-23	38	99	(AE009962) A. fulgidus predicted coding region AF2041 [Archaeoglobus fulgidus]
435	Bt1Gc688	Bt1G539	1-606	g2632226	452	506	1.80E-48	46	69	(AJ222587) YkuE protein [Bacillus subtilis]
436	Bt1Gc691	Bt1G540	1-636	g2634067	319	349	7.90E-32	37	54	(Z99112) penicillin-binding protein [Bacillus subtilis]
436	Bt1Gc691	Bt1G541	668-1043	g3183454	381	401	2.40E-37	60	43	HYPOTHETICAL 30.7 KD PROTEIN IN MCPC-KINA INTERGENIC REGION [Bacillus subtilis]
437	Bt1Gc690	Bt1G542	1-256	g3023490	245	259	2.70E-22	55	35	CR(VI) REDUCTASE [Pseudomonas sp.]
438	Bt1Gc689	Bt1G543	576-1	g3183496	288	267	3.90E-23	35	92	HYPOTHETICAL 23.4 KD PROTEIN IN AAPA-SIGV INTERGENIC REGION [Bacillus subtilis]
439	Bt1Gc692	Bt1G544	1061-1	g1769558	787	842	4.50E-84	44	75	(U81184) phospho-beta-glucosidase [Fusobacterium mortiferum]
439	Bt1Gc692	Bt1G545	1320-1259	g130298	86	102	6.70E-05	95	6	1-PHOSPHATIDYLINOSITOL PHOSPHODIESTERASE PRECURSOR (PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE C) (PI-PLC) [Bacillus thuringiensis]
440	Bt1Gc693	Bt1G546	1-759	g1731065	605	574	1.10E-55	49	78	HYPOTHETICAL 37.1 KD PROTEIN IN BMRU-ANSR INTERGENIC REGION [Bacillus subtilis]
440	Bt1Gc693	Bt1G547	779-1200	g1731066	406	417	4.90E-39	59	79	HYPOTHETICAL 19.7 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
441	Bt1Gc694	Bt1G548	32-885	g2635731	697	767	4.00E-76	52	82	(Z99120) alternate gene name: yutA [Bacillus subtilis]
442	Bt1Gc698	Bt1G549	237-1332	g3123253	1348	1261	1.80E-128	72	69	PHOSPHOENOLPYRUVATE CARBOXYKINASE [Bacillus subtilis]
443	Bt1Gc697	Bt1G550	1-911	g2619019	806	512	4.20E-49	54	95	(AF027868) putative transporter [Bacillus subtilis]
443	Bt1Gc697	Bt1G551	930-1263	g281486	152	169	9.40E-13	33	93	urf120 - Paracoccus denitrificans []
444	Bt1Gc695	Bt1G552	506-1	g134649	753	753	1.20E-74	79	82	SUPEROXIDE DISMUTASE (MN) [Bacillus caldotenax]
445	Bt1Gc701	Bt1G553	642-140	g417115	319	344	2.70E-31	37	36	PROTOPORPHYRINOGEN OXIDASE (PPO) [Bacillus subtilis]
446	Bt1Gc700	Bt1G554	1-343	g2497392	565	486	2.40E-46	97	46	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
446	Bt1Gc700	Bt1G555	1556-1342	g2984723	117	138	1.80E-09	49	99	(AF053927) probable spore germination protein F [Bacillus cereus]
447	Bt1Gc703	Bt1G556	461-147	g1787091	271	291	1.10E-25	54	99	(AE000188) orf, hypothetical protein [Escherichia coli]
448	Bt1Gc704	Bt1G557	1016-69	g2619050	280	313	5.20E-28	49	54	(AF027868) YocH [Bacillus subtilis]
449	Bt1Gc705	Bt1G558	199-1	g1731007	150	152	5.90E-11	40	79	HYPOTHETICAL 9.7 KD PROTEIN IN CCCA-SODA INTERGENIC REGION [Bacillus subtilis]
449	Bt1Gc705	Bt1G559	346-914	g1731008	654	508	1.10E-48	68	65	HYPOTHETICAL 32.5 KD PROTEIN IN CCCA-SODA INTERGENIC REGION [Bacillus subtilis]
450	Bt1Gc708	Bt1G560	427-1	g417115	257	278	7.40E-24	41	30	PROTOPORPHYRINOGEN OXIDASE (PPO) [Bacillus subtilis]
451	Bt1Gc706	Bt1G561	839-255	g2072437	131	175	2.20E-13	30	99	(U93364) EpsX [Lactococcus lactis cremoris]
452	Bt1Gc712	Bt1G562	538-1063	g136321	169	230	1.80E-18	30	37	ANTHRANILATE SYNTHASE COMPONENT I [Clostridium thermocellum]
453	Bt1Gc707	Bt1G563	204-938	g2619056	813	638	1.90E-62	63	55	(AF027868) putative transporter [Bacillus subtilis]
454	Bt1Gc711	Bt1G564	1369-487	g729343	627	641	9.00E-63	44	73	PROTEIN DLT-D PRECURSOR [Bacillus subtilis]
455	Bt1Gc710	Bt1G565	1-588	g2635798	335	347	1.30E-31	37	86	(Z99120) similar to two-component response regulator [Bacillus subtilis]
456	Bt1Gc714	Bt1G566	99-738	g2337813	696	500	7.90E-48	66	38	(Y13937) YloV protein [Bacillus subtilis]
457	Bt1Gc717	Bt1G569	1-950	g2633168	754	728	5.50E-72	48	98	(Z99108) similar to iron(III) citrate transport permease [Bacillus subtilis]
457	Bt1Gc717	Bt1G570	1097-1290	g1706292	119	162	9.80E-12	44	22	D-ALANINE AMINOTRANSFERASE (D-ASPARTATE AMINOTRANSFERASE) (D-AMINO ACID AMINOTRANSFERASE) (D-AMINO ACID TRANSAMINASE) [Bacillus licheniformis]
458	Bt1Gc709	Bt1G571	413-1110	g3915501	770	681	5.20E-67	61	96	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
										IN SPOIIIE-PCSA INTERGENIC REGION [Bacillus subtilis]
459	Bt1Gc716	Bt1G572	428-1	g2635910	101	151	4.10E-10	29	38	(Z99121) transcriptional regulator (LacI family) [Bacillus subtilis]
459	Bt1Gc716	Bt1G573	1497-479	g2117451	413	507	1.40E-48	35	100	homoserine dehydrogenase (EC 1.1.1.3) - Methanococcus jannaschii [Methanococcus jannaschii]
460	Bt1Gc719	Bt1G574	1-579	g4512389	350	259	2.60E-21	38	29	(AB011838) methyl-accepting chemotaxis protein [Bacillus halodurans]
461	Bt1Gc721	Bt1G575	582-1	g732328	461	478	1.70E-45	45	92	HYPOTHETICAL 24.6 KD PROTEIN IN DAE-TYRZ INTERGENIC REGION [Bacillus subtilis]
462	Bt1Gc720	Bt1G576	507-1	g466196	389	399	4.00E-37	50	58	HYPOTHETICAL 31.3 KD PROTEIN IN LYSA-PPIB INTERGENIC REGION (ORFX19) [Bacillus subtilis]
462	Bt1Gc720	Bt1G577	692-1123	g461900	556	500	7.90E-48	73	99	PEPTIDYL-PROLYL CIS-TRANS ISOMERASE B (PPIASE B) (ROTAMASE B) [Bacillus subtilis]
463	Bt1Gc722	Bt1G578	1-1277	g1856977	1383	1088	3.90E-110	63	83	(D50098) multidrug transporter [Bacillus subtilis]
464	Bt1Gc723	Bt1G579	275-1024	g2633809	656	685	2.00E-67	53	100	(Z99111) transcriptional regulator (DeoR family) [Bacillus subtilis]
464	Bt1Gc723	Bt1G580	1024-1370	g2633810	276	238	4.60E-20	53	39	(Z99111) fructose-1-phosphate kinase [Bacillus subtilis]
465	Bt1Gc725	Bt1G581	866-1036	g2226190	94	111	1.30E-06	40	41	(Y14081) hypothetical protein [Bacillus subtilis]
466	Bt1Gc726	Bt1G582	1-929	g2127339	680	692	8.80E-68	42	30	collagenase - Clostridium histolyticum [Clostridium histolyticum]
467	Bt1Gc728	Bt1G583	770-246	g4512387	213	246	6.50E-21	30	67	(AB011838) similar to B.subtilis ydhQ gene(53%-identity) [Bacillus halodurans]
468	Bt1Gc729	Bt1G584	1-185	g2612886	160	179	8.20E-14	56	48	(AF015825) unknown [Bacillus subtilis]
469	Bt1Gc731	Bt1G585	869-1	g2619029	271	344	2.70E-31	29	92	(AF027868) YoaR [Bacillus subtilis]
470	Bt1Gc732	Bt1G586	1198-1	g2094843	737	603	9.60E-59	40	87	(Z95324) mgE [Mycobacterium tuberculosis]
471	Bt1Gc733	Bt1G587	338-926	g1176955	403	420	2.40E-39	42	49	HYPOTHETICAL 45.8 KD PROTEIN IN ACDA-NARI INTERGENIC REGION [Bacillus subtilis]
472	Bt1Gc730	Bt1G588	10-1018	g2635715	813	703	2.40E-69	50	84	(Z99120) similar to Na <sup>+</sup> /nucleoside cotransporter [Bacillus subtilis]
473	Bt1Gc734	Bt1G589	723-142	g3298508	692	472	7.30E-45	65	99	(AB015670) A2-5a orf2; hypothetical protein [Bacillus sp.]
473	Bt1Gc734	Bt1G590	836-1381	g3318590	729	729	4.30E-72	75	98	(AB015670) A2-5a orf21; hypothetical protein [Bacillus sp.]
474	Bt1Gc736	Bt1G591	1100-1	g3599372	313	305	9.40E-27	23	73	(AF082668) CsrS [Streptococcus pyogenes]
475	Bt1Gc718	Bt1G592	687-1	g4007698	464	493	4.40E-47	44	63	(AL034446) putative integral membrane nucleotide binding

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
476	Bt1Gc738	Bt1G593	804-1	g118908	451	561	2.70E-54	42	35	protein [Streptomyces coelicolor] XAA-PRO DIPEPTIDYL-PEPTIDASE (X-PRO DIPEPTIDYL-PEPTIDASE) (X-PROLYL-DIPEPTIDYL AMINOPEPTIDASE) (X-PDAP) [Lactococcus lactis]
477	Bt1Gc740	Bt1G594	183-843	g2633511	318	375	1.40E-34	40	99	(Z99110) similar to lytic transglycosylase [Bacillus subtilis]
478	Bt1Gc742	Bt1G595	1020-1	g2340008	785	779	2.10E-77	47	97	(Z98682) YlbL protein [Bacillus subtilis]
478	Bt1Gc742	Bt1G596	1372-1025	g2340007	308	310	1.10E-27	51	44	(Z98682) YlbK protein [Bacillus subtilis]
479	Bt1Gc741	Bt1G597	1-1101	g2507017	1829	1732	2.20E-178	98	98	HEMOLYSIN BL BINDING COMPONENT PRECURSOR (ENTEROTOXIN 40 KD SUBUNIT) [Bacillus cereus]
480	Bt1Gc744	Bt1G598	1-1086	g2196513	187	302	7.60E-27	31	68	(U77778) putative membrane protein [Staphylococcus epidermidis]
480	Bt1Gc744	Bt1G599	1137-1278	g2271389	138	154	1.90E-10	60	12	(AF009352) ATPase [Bacillus subtilis]
481	Bt1Gc746	Bt1G600	1-643	g2636487	867	804	4.80E-80	80	54	(Z99124) pyrimidine-nucleoside transport protein [Bacillus subtilis]
481	Bt1Gc746	Bt1G601	683-979	g730287	378	413	1.30E-38	80	23	PYRIMIDINE-NUCLEOSIDE PHOSPHORYLASE [Bacillus subtilis]
482	Bt1Gc739	Bt1G602	102-587	g3238042	69	133	6.10E-09	24	99	(AP000006) 166aa long hypothetical protein [Pyrococcus horikoshii]
483	Bt1Gc750	Bt1G603	1-973	g1856977	536	506	1.80E-48	35	63	(D50098) multidrug transporter [Bacillus subtilis]
484	Bt1Gc751	Bt1G604	637-41	g732240	338	395	1.10E-36	40	99	HYPOTHETICAL PROTEIN IN PEPC 5'REGION (ORF 2) [Lactococcus lactis]
484	Bt1Gc751	Bt1G605	844-1209	g4104605	377	390	3.60E-36	63	54	(AF036967) putative response regulator [Lactobacillus sakei]
485	Bt1Gc752	Bt1G606	260-1	g129799	173	190	3.80E-14	41	18	AMINOACYL-HISTIDINE DIPEPTIDASE (XAA-HIS DIPEPTIDASE) (X-HIS DIPEPTIDASE) (BETA- ALANYL-HISTIDINE DIPEPTIDASE) (CARNOSINASE) (PEPTIDASE D) [Escherichia coli]
486	Bt1Gc748	Bt1G607	1-302	g1651667	195	222	2.30E-18	43	31	(D90899) regulatory protein PchR [Synecocystis sp.]
486	Bt1Gc748	Bt1G608	422-1100	g3116222	407	388	5.80E-36	34	57	(AB007122) transporter [Arthrobacter sp.]
487	Bt1Gc753	Bt1G609	1310-1	g1381682	928	993	4.50E-100	45	97	(U58864) orf1 [Bacillus subtilis]
488	Bt1Gc754	Bt1G610	1-275	g1176283	169	172	4.50E-13	36	56	HYPOTHETICAL 18.8 KD PROTEIN IN GNTR-GGT INTERGENIC REGION (O162) [Escherichia coli]
489	Bt1Gc755	Bt1G611	945-293	g729418	908	853	3.10E-85	77	100	PROBABLE ENDONUCLEASE III (DNA-(APURINIC OR APYRIMIDINIC SITE) LYASE) [Bacillus subtilis]
490	Bt1Gc747	Bt1G612	1-819	g418587	741	774	7.30E-77	56	84	HYPOTHETICAL 34.0 KD PROTEIN IN RHO-MURA

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
491	Bt1Gc756	Bt1G613	1-714	g3121979	459	494	3.40E-47	43	84	INTERGENIC REGION (ORFQ) [Bacillus subtilis] D-ALANINE AMINOTRANSFERASE (D-ASPARTATE AMINOTRANSFERASE) (D-AMINO ACID AMINOTRANSFERASE) (D-AMINO ACID TRANSAMINASE) [Bacillus subtilis]
491	Bt1Gc756	Bt1G614	813-1242	g2499210	84	135	3.80E-09	40	32	PUTATIVE NAD(P)H NITROREDUCTASE YDGI [Bacillus subtilis]
492	Bt1Gc757	Bt1G615	1-1039	g2982791	287	359	6.90E-33	41	44	(AE000670) hypothetical protein [Aquifex aeolicus]
493	Bt1Gc759	Bt1G616	608-1	g729175	823	737	6.10E-73	79	78	CODY PROTEIN (VEGETATIVE PROTEIN 286B) (VEG286B) [Bacillus subtilis]
493	Bt1Gc759	Bt1G617	1229-689	g2126915	754	770	1.90E-76	82	39	codX protein - Bacillus subtilis []
494	Bt1Gc760	Bt1G618	945-131	g4514323	779	820	9.70E-82	53	74	(AB013367) unknown [Bacillus halodurans]
495	Bt1Gc761	Bt1G619	895-1	g114116	899	856	1.50E-85	57	62	ARGININE/ORNITHINE ANTIPORTER [Pseudomonas aeruginosa]
495	Bt1Gc761	Bt1G620	1817-1127	g4155319	400	441	1.40E-41	39	41	(AE001506) putative osmoprotection binding protein [Helicobacter pylori J99]
496	Bt1Gc763	Bt1G621	1018-1	g1805447	254	322	5.70E-29	25	93	(D50453) response-regulator aspartate phosphatase C [Bacillus subtilis]
496	Bt1Gc763	Bt1G622	1018-1	g1648857	202	243	1.90E-20	22	93	(Z81356) aspartyl-phosphate phosphatases [Bacillus subtilis]
497	Bt1Gc766	Bt1G623	105-1	g1934629	131	131	1.00E-08	69	39	(U93875) YraL [Bacillus subtilis]
497	Bt1Gc766	Bt1G624	914-478	g2619056	401	337	1.50E-30	58	33	(AF027868) putative transporter [Bacillus subtilis]
498	Bt1Gc762	Bt1G625	468-1334	g2634335	550	587	4.80E-57	40	100	(Z99114) similar to macrolide glycosyltransferase [Bacillus subtilis]
499	Bt1Gc765	Bt1G626	270-40	g2633909	100	157	1.80E-11	40	99	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
499	Bt1Gc765	Bt1G627	1216-440	g133475	1164	1175	2.30E-119	89	100	RNA POLYMERASE SIGMA-G FACTOR (STAGE III SPORULATION PROTEIN G) [Bacillus subtilis]
499	Bt1Gc765	Bt1G628	1864-1377	g133289	755	719	4.90E-71	96	68	RNA POLYMERASE SIGMA-35 FACTOR PRECURSOR [Bacillus thuringiensis]
500	Bt1Gc767	Bt1G629	1-1260	g1865711	592	646	2.70E-63	51	78	(Y11477) endolysin [Bacteriophage Bastille]
501	Bt1Gc758	Bt1G630	1149-1	g1945050	1844	1685	2.10E-173	98	86	(U63928) L2 protein [Bacillus cereus]
502	Bt1Gc768	Bt1G631	1-816	g2117698	898	904	1.20E-90	63	88	aspartate carbamoyltransferase (EC 2.1.3.2) - Bacillus caldolyticus []
503	Bt1Gc771	Bt1G632	1421-1	g2497686	69	130	6.80E-16	24	50	PLATELET-ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (PAF)

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
										ACETYLHYDROLASE) (PAF 2-ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL- PLA(2)) (2-ACETYL-1- ALKYLGLYCEROPHOSPHOCOLINE ESTERASE) (1- ALKYL-2-ACETYLGLYCEROPHO... [Cavia porcellus] OLIGOPEPTIDE TRANSPORT ATP-BINDING PROTEIN OPPF [Bacillus subtilis]
504	Bt1Gc769	Bt1G633	291-1	g129188	369	369	6.00E-34	72	31	(Z99110) oligopeptide ABC transporter (ATP-binding protein) [Bacillus subtilis]
504	Bt1Gc769	Bt1G634	775-266	g2633500	597	615	5.10E-60	67	47	(AB001488) PROBABLE TRANSPORT PROTEIN, SIMILAR PRODUCT IN MYCOBACTERIUM LEPRAE. [Bacillus subtilis]
505	Bt1Gc774	Bt1G635	27-1037	g1881371	284	300	1.80E-25	33	32	9 ORF IS231C [Bacillus thuringiensis] 84 (AP000007) 306aa long hypothetical UDP-glucose 4- epimerase [Pyrococcus horikoshii] 49 TRANSKETOLASE [Bacillus subtilis] 20 ORF IS231C [Bacillus thuringiensis] 62 (AF021937) catalase 1 [Arabidopsis thaliana] 98 kinB protein - Bacillus subtilis [] 99 HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN PRKA-CSPB INTERGENIC REGION [Bacillus subtilis]
505	Bt1Gc774	Bt1G636	1037-904	g225559	184	202	1.80E-15	91	9	(AJ005075) enzyme I [Bacillus megaterium]
506	Bt1Gc773	Bt1G637	241-1020	g3258173	217	298	2.00E-26	30	99	(AF065404) pXO1-73 [Bacillus anthracis] (Z75208) hypothetical protein [Bacillus subtilis] 99 HYPOTHETICAL 22.0 KD PROTEIN IN RECQ-CMK INTERGENIC REGION [Bacillus subtilis] 93 DEOXYRIBOSE-PHOSPHATE ALDOLASE (PHOSPHODEOXYRIBOALDOLASE) (DEOXYRIBOALDOLASE) [Bacillus subtilis]
507	Bt1Gc776	Bt1G638	987-1	g2507484	1171	1103	1.00E-111	65	46	PYRIMIDINE NUCLEOSIDE TRANSPORT PROTEIN [Bacillus subtilis]
508	Bt1Gc779	Bt1G639	1294-1011	g225559	417	447	3.30E-42	93	100	(Y09323) hypothetical protein [Bacillus cereus]
509	Bt1Gc775	Bt1G640	1-1298	g2511725	627	687	1.20E-67	42	76	(U66480) YnaE [Bacillus subtilis]
510	Bt1Gc778	Bt1G641	1976-757	g421510	447	473	5.70E-45	32	28	SHORT-CHAIN FATTY ACIDS TRANSPORTER
511	Bt1Gc780	Bt1G642	965-1333	g1724000	121	131	1.00E-08	31		
512	Bt1Gc782	Bt1G643	1216-1	g3021327	1603	1509	9.50E-155	77	71	
513	Bt1Gc788	Bt1G644	532-266	g4894289	112	112	1.00E-06	27	99	
513	Bt1Gc788	Bt1G645	1085-634	g1770032	408	434	7.80E-41	58	48	
514	Bt1Gc784	Bt1G646	42-900	g1730886	387	349	7.90E-32	43	99	
515	Bt1Gc786	Bt1G647	1-593	g1706363	745	660	8.80E-65	74	93	
515	Bt1Gc786	Bt1G648	735-1277	g1709415	685	558	5.60E-54	72	46	
516	Bt1Gc785	Bt1G649	989-350	g2073393	596	617	3.20E-60	55	100	
516	Bt1Gc785	Bt1G650	1524-1031	g1750116	245	319	1.20E-28	38	76	
517	Bt1Gc783	Bt1G651	1211-1588	g1175305	285	295	5.90E-26	41	28	



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
518	Bt1Gc790	Bt1G652	1-306	g2634340	302	307	2.20E-27	54	87	[Haemophilus influenzae Rd]
518	Bt1Gc790	Bt1G653	223-723	g3256904	104	175	2.20E-13	26	60	(Z99114) yojF [Bacillus subtilis]
518	Bt1Gc790	Bt1G654	559-723	g2634339	197	197	1.00E-15	65	38	(AP000002) 272aa long hypothetical protein [Pyrococcus horikoshii]
519	Bt1Gc791	Bt1G655	1351-125	g137192	1506	1530	5.60E-157	72	67	(Z99114) yojG [Bacillus subtilis]
520	Bt1Gc793	Bt1G656	705-1	g2313187	359	380	8.00E-35	37	40	EXCINUCLEASE ABC SUBUNIT C [Bacillus subtilis]
521	Bt1Gc794	Bt1G657	1063-112	g2088525	432	450	1.60E-42	35	100	(AE000532) 2',3'-cyclic-nucleotide 2'-phosphodiesterase (cpdB) [Helicobacter pylori 26695]
522	Bt1Gc798	Bt1G658	557-78	g1731124	658	512	4.20E-49	76	99	(U70619) heroin esterase [Rhodococcus sp.]
522	Bt1Gc798	Bt1G659	824-621	g2635515	159	168	1.20E-12	43	99	HYPOTHETICAL 17.8 KD PROTEIN IN BLTR-SPOIIC INTERGENIC REGION [Bacillus subtilis]
523	Bt1Gc795	Bt1G660	1-1737	g2501051	778	869	6.20E-87	33	98	(Z99119) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
524	Bt1Gc797	Bt1G661	1440-753	g1176993	633	645	3.40E-63	53	100	PROBABLE ARGINYL-TRNA SYNTHETASE, CYTOPLASMIC (ARGININE--TRNA LIGASE) (ARGRS) [Saccharomyces cerevisiae]
525	Bt1Gc796	Bt1G662	1-487	g1825622	97	118	2.30E-08	25	44	HYPOTHETICAL 26.6 KD SENSORY TRANSDUCTION PROTEIN IN IDH 3'REGION [Bacillus subtilis]
526	Bt1Gc799	Bt1G663	684-1	g2634058	918	924	9.30E-93	75	55	(U88170) No definition line found [Caenorhabditis elegans]
526	Bt1Gc799	Bt1G664	1307-787	g2634057	498	476	2.80E-45	58	74	(Z99112) similar to processing protease [Bacillus subtilis]
527	Bt1Gc800	Bt1G665	957-1	g2127359	169	209	5.40E-17	53	29	(Z99112) similar to processing protease [Bacillus subtilis]
528	Bt1Gc802	Bt1G666	1-1178	g129021	1524	1481	8.80E-152	73	92	hypothetical protein 3 - Clostridium perfringens [Clostridium perfringens]
529	Bt1Gc803	Bt1G667	1-1083	g1352313	597	635	3.90E-62	35	66	SPOOB-ASSOCIATED GTP-BINDING PROTEIN [Bacillus subtilis]
530	Bt1Gc805	Bt1G668	497-1366	g131027	448	436	4.80E-41	39	100	DIPEPTIDE-BINDING PROTEIN DPPE PRECURSOR [Bacillus subtilis]
531	Bt1Gc806	Bt1G669	74-1	g1146349	90	106	4.50E-06	92	38	PROTEIN EXPORT PROTEIN PRSA PRECURSOR [Bacillus subtilis]
531	Bt1Gc806	Bt1G670	1331-183	g1945051	1875	1135	9.10E-158	97	100	(L43071) ORF [Edwardsiella tarda]
532	Bt1Gc807	Bt1G671	1-995	g2982764	495	477	2.20E-45	36	85	(U63928) L1 protein [Bacillus cereus]
533	Bt1Gc808	Bt1G672	718-191	g2635678	550	562	2.10E-54	59	99	(AE000669) N-acetylornithine aminotransferase [Aquifex aeolicus]
533	Bt1Gc808	Bt1G673	1104-892	g1770028	260	266	4.90E-23	73	99	(Z99120) yueE [Bacillus subtilis]
										(Z75208) hypothetical protein [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
533	Bt1Gc808	Bt1G674	1220-1345	g1770029	133	133	1.20E-08	62	17	(Z75208) hypothetical protein [Bacillus subtilis]
534	Bt1Gc809	Bt1G675	1-872	g1709639	1110	1110	1.80E-112	71	71	PEPTIDASE T (AMINOTRIPEPTIDASE) (TRIPEPTIDASE) [Bacillus subtilis]
534	Bt1Gc809	Bt1G676	1828-917	g477274	836	794	5.50E-79	52	100	3-methyladenine DNA glycosylase - Bacillus subtilis []
535	Bt1Gc804	Bt1G677	1605-1	g3322922	340	453	9.80E-42	27	41	(AE001237) exonuclease (sbcC) [Treponema pallidum]
536	Bt1Gc812	Bt1G678	273-1253	g1730898	1225	1112	1.10E-112	75	100	HYPOTHETICAL 36.3 KD PROTEIN IN RECQ-CMK INTERGENIC REGION [Bacillus subtilis]
537	Bt1Gc810	Bt1G679	914-1	g2506131	1046	1031	4.20E-104	69	34	ACONITATE HYDRATASE (CITRATE HYDRO-LYASE) (ACONITASE) [Bacillus subtilis]
538	Bt1Gc811	Bt1G680	747-1	g98515	1116	1016	1.70E-102	93	90	flagellin B chain - Bacillus thuringiensis [Bacillus thuringiensis alesti]
539	Bt1Gc814	Bt1G681	1-859	g1881339	261	108	6.60E-15	67	48	(AB001488) PROBABLE TRANSCRIPTIONAL REGULATOR, SIMILAR TO REGULATOR OF ANTIBIOTIC TRANSPORT COMPLEXES IN STREPTOMYCES HYGROSCOPICUS. [Bacillus subtilis]
539	Bt1Gc814	Bt1G682	1829-1	g1169000	262	353	5.60E-31	36	19	MICROBIAL COLLAGENASE PRECURSOR (120 KD COLLAGENASE) [Clostridium perfringens]
540	Bt1Gc813	Bt1G683	525-1092	g96807	269	359	8.70E-32	41	21	Mg2+-transporting ATPase (EC 3.6.1.-) mgfB - Salmonella typhimurium [Salmonella typhimurium]
541	Bt1Gc817	Bt1G684	875-1	g1731017	645	562	2.10E-54	48	40	HYPOTHETICAL 80.1 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
542	Bt1Gc815	Bt1G685	1303-117	g2635180	1637	1597	4.50E-164	78	93	(Z99117) similar to protease [Bacillus subtilis]
543	Bt1Gc816	Bt1G686	1164-728	g1708641	205	242	1.70E-20	33	46	(U51115) YeaC [Bacillus subtilis]
544	Bt1Gc818	Bt1G687	102-485	g2635382	283	301	9.70E-27	43	99	(Z99118) similar to hypothetical proteins [Bacillus subtilis]
545	Bt1Gc819	Bt1G688	1-142	g2633732	140	161	6.60E-12	58	22	(Z99111) similar to hypothetical proteins [Bacillus subtilis]
545	Bt1Gc819	Bt1G689	141-671	g2633733	643	659	1.10E-64	69	99	(Z99111) similar to hypothetical proteins [Bacillus subtilis]
545	Bt1Gc819	Bt1G690	1963-1612	g3183454	353	388	5.80E-36	63	41	HYPOTHETICAL 30.7 KD PROTEIN IN MCPC-KINA INTERGENIC REGION [Bacillus subtilis]
546	Bt1Gc820	Bt1G691	848-101	g2497382	606	641	9.00E-63	94	31	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
546	Bt1Gc820	Bt1G692	565-848	g225559	447	461	1.10E-43	97	20	ORF IS231C [Bacillus thuringiensis]
547	Bt1Gc821	Bt1G693	198-1	g1434920	107	101	0.00031	39	8	(D86346) crystal protein [Bacillus thuringiensis]
547	Bt1Gc821	Bt1G694	951-585	g2226238	452	484	3.90E-46	72	35	(Y14083) hypothetical protein [Bacillus subtilis]
548	Bt1Gc823	Bt1G695	581-234	g2522004	464	467	2.50E-44	70	99	(AB007637) function unknown [Bacillus subtilis]
549	Bt1Gc824	Bt1G696	1-1465	g1255196	1768	1787	3.30E-184	64	83	(U50744) BSMA [Bacillus stearothermophilus]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat- nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
550	Bt1Gc825	Bt1G697	1-1074	g2635600	899	652	6.20E-64	48	93	(Z99119) similar to hypothetical proteins [Bacillus subtilis]
551	Bt1Gc822	Bt1G698	854-1	g2558482	1125	1130	1.40E-114	74	72	(AB008120) phosphopentomutase [Bacillus stearothermophilus]
552	Bt1Gc826	Bt1G699	1-800	g4835313	127	185	1.90E-14	33	88	(AL049863) hypothetical protein [Streptomyces coelicolor]
552	Bt1Gc826	Bt1G700	687-1455	g1644210	687	679	8.50E-67	53	100	(D64126) unknown [Bacillus subtilis]
553	Bt1Gc829	Bt1G702	274-1	g118797	168	182	1.50E-13	39	24	DNA POLYMERASE III, BETA CHAIN [Bacillus subtilis]
553	Bt1Gc829	Bt1G703	712-351	g2129134	92	156	2.20E-11	28	72	mutator protein mutT - Methanococcus jannaschii [Methanococcus jannaschii]
553	Bt1Gc829	Bt1G704	1459-1034	g1181177	195	220	3.70E-18	30	99	(L25828) ORF 12 [Erwinia amylovora]
554	Bt1Gc830	Bt1G705	1-149	g2274866	124	143	5.30E-10	53	71	(AJ000005) orf1 [Bacillus megaterium]
554	Bt1Gc830	Bt1G706	158-972	g2274867	997	716	1.00E-70	69	84	(AJ000005) glucose kinase [Bacillus megaterium]
555	Bt1Gc833	Bt1G707	1083-878	g3005554	86	113	4.00E-06	29	22	(AF047044) putative transposase [Anabaena PCC7120]
556	Bt1Gc828	Bt1G708	1929-420	g728913	1323	1234	1.30E-125	51	56	MG(2+) TRANSPORT ATPASE, P-TYPE 1 [Escherichia coli]
557	Bt1Gc834	Bt1G709	1110-1	g1730904	1129	1108	2.90E-112	59	82	HYPOTHETICAL 51.2 KD PROTEIN IN RECQ-CMK INTERGENIC REGION [Bacillus subtilis]
557	Bt1Gc834	Bt1G710	1639-1129	g1644192	833	771	1.50E-76	94	66	(D63645) spore cortex-lytic enzyme prepeptide [Bacillus cereus]
558	Bt1Gc835	Bt1G711	41-598	g2226257	432	131	2.30E-20	45	99	(Y14084) hypothetical protein [Bacillus subtilis]
558	Bt1Gc835	Bt1G712	686-1308	g2226256	344	418	3.90E-39	42	43	(Y14084) hypothetical protein [Bacillus subtilis]
559	Bt1Gc836	Bt1G713	643-251	g629241	134	212	2.60E-17	34	75	lmbK protein - Streptomyces lincolnensis [Streptomyces lincolnensis]
559	Bt1Gc836	Bt1G714	1162-1	g3915419	47	124	5.50E-08	36	56	HYPOTHETICAL 17.9 KD PROTEIN IN PHOB-GROES INTERGENIC REGION [Bacillus subtilis]
560	Bt1Gc839	Bt1G715	256-1	g2634117	264	287	2.90E-25	60	27	(Z99113) tRNA isopentenylpyrophosphate transferase [Bacillus subtilis]
560	Bt1Gc839	Bt1G716	1152-751	g2619050	277	319	1.20E-28	52	45	(AF027868) Yoch [Bacillus subtilis]
561	Bt1Gc838	Bt1G717	1246-468	g586817	998	999	1.00E-100	77	84	HYPOTHETICAL 34.0 KD PROTEIN IN COTF-TETB INTERGENIC REGION [Bacillus subtilis]
562	Bt1Gc840	Bt1G718	1106-203	g3355675	261	363	2.60E-33	32	89	(AL031124) hypothetical protein SC1C2.08 [Streptomyces coelicolor]
563	Bt1Gc841	Bt1G719	488-710	g4874247	80	138	3.30E-09	34	29	(AF041061) beta-lactamase-like protein [Mycobacterium smegmatis]
564	Bt1Gc842	Bt1G720	198-1391	g585250	131	218	6.00E-18	28	90	HETI PROTEIN II
565	Bt1Gc843	Bt1G721	1451-315	g1731316	1136	1157	1.90E-117	55	100	HYPOTHETICAL 41.6 KD PROTEIN IN IDH-DEOR

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
566	Bt1Gc846	Bt1G722	1-1371	g2619033	1606	1624	6.20E-167	69	52	INTERGENIC REGION [Bacillus subtilis]
567	Bt1Gc850	Bt1G723	1-416	g113673	269	349	7.90E-32	53	53	(AF027868) PEP synthase [Bacillus subtilis]
568	Bt1Gc844	Bt1G724	1-381	g132678	541	486	2.40E-46	84	86	N-ACETYLMURAMOYL-L-ALANINE AMIDASE
568	Bt1Gc844	Bt1G725	384-1333	g421457	1153	990	9.40E-100	72	73	CWLA PRECURSOR (CELL WALL HYDROLASE)
569	Bt1Gc847	Bt1G726	710-1	g1742360	201	253	1.20E-21	32	74	(AUTOLYSIN) [Bacillus sp.]
570	Bt1Gc832	Bt1G727	1152-619	g1176958	132	80	3.60E-08	29	99	50S RIBOSOMAL PROTEIN L15 []
571	Bt1Gc845	Bt1G728	1-1178	g1731015	1139	955	4.80E-96	56	91	secY protein - Bacillus licheniformis []
572	Bt1Gc854	Bt1G729	1-424	g2633836	327	339	9.10E-31	45	67	(D90784) Phosphothricin acetyltransferase (EC 2.3.1.-).
572	Bt1Gc854	Bt1G730	766-503	g2633835	260	263	1.00E-22	55	99	[Escherichia coli]
572	Bt1Gc854	Bt1G731	1134-1562	g118334	537	530	5.20E-51	74	29	INTERGENIC REGION [Bacillus subtilis]
573	Bt1Gc852	Bt1G732	1263-522	g1865705	281	327	1.70E-29	36	100	LYSINE DECARBOXYLASE (LDC) [Bacillus subtilis]
574	Bt1Gc857	Bt1G733	10-717	g2636560	390	403	1.50E-37	37	100	(Y11476) N-acetylmuramoyl-L-alanine amidase
574	Bt1Gc857	Bt1G734	1309-1477	g1881344	103	128	1.90E-07	40	12	[Bacteriophage 12826]
575	Bt1Gc853	Bt1G735	1879-1	g2226165	1890	1770	2.10E-182	59	94	(Z99124) similar to transcriptional regulator (GntR family)
576	Bt1Gc855	Bt1G737	393-1	g4033506	582	583	1.30E-56	88	45	[Bacillus subtilis]
576	Bt1Gc855	Bt1G738	972-607	g745617	594	594	8.60E-58	98	99	(AB001488) SIMILAR TO THE RHIZOPINE
576	Bt1Gc855	Bt1G739	1174-1537	g139993	206	240	2.80E-20	45	68	CATABOLISM (MOCR) GENE OF RHIZOBIUM
577	Bt1Gc858	Bt1G740	871-1	g2293165	803	792	9.00E-79	54	46	MELILOT. [Bacillus subtilis]
577	Bt1Gc858	Bt1G741	1239-1567	g3258280	112	139	1.40E-09	39	63	(Y14080) hypothetical protein [Bacillus subtilis]
578	Bt1Gc863	Bt1G742	354-1	g2340007	345	345	2.10E-31	54	45	FRUCTOSE-BISPHOSPHATE ALDOLASE [Bacillus stearothermophilus]
578	Bt1Gc863	Bt1G743	465-1486	g2340006	996	1007	1.50E-101	58	83	spoOF gene [Bacillus thuringiensis]
579	Bt1Gc865	Bt1G744	1507-326	g1684651	638	349	9.00E-55	33	100	HYPOTHETICAL 19.1 KD PROTEIN IN SPOOF-PYRG
										INTERGENIC REGION (ORFS) [Bacillus subtilis]
										(AF008220) asparagine synthase [Bacillus subtilis]
										(AP000007) 172aa long hypothetical protein [Pyrococcus horikoshii]
										(Z98682) YlbK protein [Bacillus subtilis]
										(Z98682) YlbJ protein [Bacillus subtilis]
										(Z82987) unknown similar to quinolon resistance protein

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
580	Bt1Gc860	Bt1G745	1269-263	g1711644	1120	1112	1.10E-112	63	65	NorA [Bacillus subtilis] 65 PROBABLE ASPARAGINYL-TRNA SYNTHETASE (ASPARAGINE--TRNA LIGASE) (ASNRS) [Synechocystis sp.]
581	Bt1Gc866	Bt1G746	366-1224	g2634335	570	606	4.60E-59	42	100	(Z99114) similar to macrolide glycosyltransferase [Bacillus subtilis]
582	Bt1Gc867	Bt1G747	315-756	g2632035	342	362	3.30E-33	50	99	(AJ002571) YkmA [Bacillus subtilis]
582	Bt1Gc867	Bt1G748	780-1190	g2632036	319	368	7.70E-34	51	99	(AJ002571) YknA [Bacillus subtilis]
582	Bt1Gc867	Bt1G749	1490-1239	g586754	158	195	1.60E-15	45	43	HYPOTHETICAL GTP-BINDING PROTEIN IN LONA- HEMA INTERGENIC REGION (OREX) [Bacillus subtilis]
583	Bt1Gc861	Bt1G750	1437-1	g416605	1013	1072	1.90E-108	42	88	MEDIUM-CHAIN-FATTY-ACID--COA LIGASE (MEDIUM-CHAIN ACYL-COA SYNTHETASE) [Pseudomonas oleovorans]
584	Bt1Gc868	Bt1G751	495-1	g2506820	57	159	3.30E-11	26	49	ACETYL ESTERASE [Escherichia coli]
584	Bt1Gc868	Bt1G752	1486-1108	g2632453	289	267	3.90E-23	52	71	(Z99104) ybcF [Bacillus subtilis]
585	Bt1Gc869	Bt1G753	1-968	g131493	1182	1079	3.50E-109	73	46	PTS SYSTEM, GLUCOSE-SPECIFIC IIABC COMPONENT (EIIABC-GLC) (GLUCOSE-PERMEASE IIABC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, ABC COMPONENT) (EII-GLC / EIII-GLC) [Bacillus subtilis]
586	Bt1Gc862	Bt1G754	1-1141	g2632042	452	524	2.30E-50	33	76	(AJ002571) YkoD [Bacillus subtilis]
587	Bt1Gc871	Bt1G755	1-353	g1197032	334	344	2.70E-31	54	94	(M29945) ORF 2; putative [Yersinia enterocolitica]
587	Bt1Gc871	Bt1G756	435-1100	g3257896	191	283	7.80E-25	37	99	(AP000006) 197aa long hypothetical protein [Pyrococcus horikoshii]
587	Bt1Gc871	Bt1G757	1414-1138	g1176701	333	351	4.90E-32	74	43	HYPOTHETICAL 23.6 KD PROTEIN IN QCRC-DAPB INTERGENIC REGION [Bacillus subtilis]
588	Bt1Gc872	Bt1G758	2061-1	g627059	89	185	8.90E-11	19	20	liver stage antigen LSA-1 - Plasmodium falciparum [Plasmodium falciparum]
588	Bt1Gc872	Bt1G759	1243-1	g3322922	230	306	5.40E-26	24	38	(AE001237) exonuclease (sbcC) [Treponema pallidum]
588	Bt1Gc872	Bt1G760	2128-1249	g3322921	100	227	1.70E-18	24	77	(AE001237) exonuclease, putative [Treponema pallidum]
589	Bt1Gc873	Bt1G761	389-1	g731171	303	216	9.80E-18	55	42	N-ACETYLMURAMOYL-L-ALANINE AMIDASE XYLA PRECURSOR (CELL WALL HYDROLASE) (AUTOLYSIN) [Bacillus subtilis]
590	Bt1Gc874	Bt1G762	475-1152	g465600	244	273	9.00E-24	30	100	HYPOTHETICAL 25.3 KD PROTEIN IN NFO-FRUA INTERGENIC REGION [Escherichia coli]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
591	Bt1Gc876	Bt1G763	1705-234	g464794	2244	2092	1.60E-216	88	100	STAGE IV SPORULATION PROTEIN A [Bacillus subtilis]
592	Bt1Gc877	Bt1G764	1-1341	g1945090	1423	1374	1.90E-140	64	97	(D88802) B. subtilis alkaline phosphatase IIIA; P19405 secretory [Bacillus subtilis]
593	Bt1Gc875	Bt1G765	1744-1	g1731300	777	696	1.30E-68	38	91	HYPOTHETICAL 70.5 KD PROTEIN IN IDH 3 REGION [Bacillus subtilis]
593	Bt1Gc875	Bt1G766	1744-512	g709994	655	519	7.70E-50	41	100	(D14399) hypothetical protein [Bacillus subtilis]
593	Bt1Gc875	Bt1G767	2285-1737	g2293177	468	505	2.30E-48	50	73	(AF008220) transporter [Bacillus subtilis]
594	Bt1Gc879	Bt1G768	216-788	g4539178	155	228	5.30E-19	28	73	(AL049485) hypothetical protein [Streptomyces coelicolor]
595	Bt1Gc880	Bt1G769	103-708	g2415385	542	481	8.10E-46	52	100	(AF015775) YodC [Bacillus subtilis]
596	Bt1Gc882	Bt1G770	699-1	g3122814	1102	1106	4.80E-112	96	61	30S RIBOSOMAL PROTEIN S1 HOMOLOG [Bacillus cereus]
596	Bt1Gc882	Bt1G771	1340-1037	g1945144	487	490	9.10E-47	98	51	(Y08953) cytidylate kinase-like protein [Bacillus cereus]
597	Bt1Gc883	Bt1G772	446-172	g2634062	223	240	2.80E-20	54	99	(Z99112) ymfL [Bacillus subtilis]
597	Bt1Gc883	Bt1G773	921-446	g1842437	566	594	8.60E-58	73	99	(U87792) unknown [Bacillus subtilis]
598	Bt1Gc881	Bt1G774	586-1003	g2982870	104	150	3.20E-10	25	45	(AE000675) NADH dehydrogenase (ubiquinone) [Aquifex aeolicus]
599	Bt1Gc884	Bt1G775	1-106	g1770012	116	132	4.30E-08	69	10	(Z75208) hypothetical protein [Bacillus subtilis]
599	Bt1Gc884	Bt1G776	752-1	g2982980	84	167	1.50E-12	31	73	(AE000682) transcriptional regulator (TetR/AcrR family) [Aquifex aeolicus]
600	Bt1Gc885	Bt1G777	541-1	g2635190	619	535	1.50E-51	66	83	(Z99117) glutamine ABC transporter (integral membrane protein) [Bacillus subtilis]
600	Bt1Gc885	Bt1G778	1155-610	g2635189	693	693	2.80E-68	73	66	(Z99117) glutamine ABC transporter (glutamine-binding protein) [Bacillus subtilis]
601	Bt1Gc888	Bt1G779	587-1	g114636	520	550	4.00E-53	53	68	ATP SYNTHASE GAMMA CHAIN [Bacillus megaterium]
602	Bt1Gc889	Bt1G780	267-1121	g3256528	241	338	1.20E-30	29	100	(AP000001) 289aa long hypothetical protein [Pyrococcus horikoshii]
602	Bt1Gc889	Bt1G781	151-1295	g1741967	152	307	2.20E-27	25	100	(D86380) Alkaline D-peptidase [Bacillus cereus]
603	Bt1Gc891	Bt1G782	1-270	g3005554	79	142	2.60E-09	29	28	(AF047044) putative transposase [Anabaena PCC7120]
604	Bt1Gc892	Bt1G783	497-1	g2116766	499	496	2.10E-47	54	65	(D86418) YfhH [Bacillus subtilis]
604	Bt1Gc892	Bt1G784	880-641	g2497392	412	412	1.70E-38	99	32	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
605	Bt1Gc890	Bt1G785	1006-179	g732410	864	642	7.10E-63	61	100	PROBABLE ABC TRANSPORTER PERMEASE PROTEIN (ORFL) [Bacillus subtilis]
605	Bt1Gc890	Bt1G786	1250-1012	g732409	136	152	2.20E-10	37	23	HYPOTHETICAL 36.3 KD LIPOPROTEIN PRECURSOR

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
606	Bt1Gc894	Bt1G787	44-972	g2633695	509	435	1.60E-43	36	100	(ORFK) [Bacillus subtilis]
607	Bt1Gc893	Bt1G788	1-1177	g2415746	1445	1437	4.00E-147	69	98	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
607	Bt1Gc893	Bt1G789	1396-2083	g1770002	189	222	2.30E-18	28	97	(AB000617) YceI [Bacillus subtilis]
608	Bt1Gc895	Bt1G790	1-946	g2633130	1329	1308	1.90E-133	79	94	(Z75208) autolysin response regulator [Bacillus subtilis]
609	Bt1Gc897	Bt1G791	1-404	g1881293	136	200	4.90E-16	35	98	(Z99108) acetoin dehydrogenase E1 component (TPP-dependent alpha subunit) [Bacillus subtilis]
610	Bt1Gc896	Bt1G792	1549-241	g3915990	954	761	1.70E-75	44	83	(AB001488) PROBABLE REPRESSOR PROTEIN. [Bacillus subtilis]
611	Bt1Gc900	Bt1G793	1-151	g3183248	102	117	3.00E-07	39	53	HYPOTHETICAL 58.3 KD PROTEIN IN GLPD-CSPB INTERGENIC REGION [Bacillus subtilis]
612	Bt1Gc901	Bt1G794	614-357	g1174408	318	137	2.30E-09	76	99	HYPOTHETICAL 10.9 KD PROTEIN IN PHRC-GDH INTERGENIC REGION [Bacillus subtilis]
613	Bt1Gc898	Bt1G795	1-2094	g2632985	1127	958	2.30E-96	37	66	STAGE V SPORULATION PROTEIN S [Bacillus subtilis]
614	Bt1Gc899	Bt1G796	748-1623	g1673402	480	530	5.20E-51	38	62	(Z99107) similar to acriflavin resistance protein [Bacillus subtilis]
615	Bt1Gc902	Bt1G797	1344-46	g3282523	436	522	3.70E-50	27	89	(Z82044) hypothetical 54.4 kd protein [Bacillus subtilis]
616	Bt1Gc903	Bt1G798	461-1	g3334798	259	282	1.00E-24	34	68	(AF025433) tyrosine/dopa decarboxylase [Papaver somniferum]
617	Bt1Gc904	Bt1G799	1306-747	g2635899	613	544	1.70E-52	60	91	(AL031107) putative two-component regulator [Streptomyces coelicolor]
618	Bt1Gc905	Bt1G800	177-908	g1162917	1026	235	6.70E-40	85	100	(Z99121) yvbH [Bacillus subtilis]
618	Bt1Gc905	Bt1G801	1001-1244	g1731004	344	351	4.90E-32	81	25	(L48554) vrrA gene product [Bacillus anthracis]
619	Bt1Gc909	Bt1G802	1318-1	g549806	1384	1407	6.10E-144	62	90	LYTB PROTEIN HOMOLOG [Bacillus subtilis]
620	Bt1Gc907	Bt1G803	1412-649	g732331	450	407	5.70E-38	42	58	HYPOTHETICAL 55.7 KD PROTEIN IN PLC 5'REGION (ORF 2) [Clostridium perfringens]
621	Bt1Gc911	Bt1G804	229-1	g1769919	118	157	3.40E-10	42	9	HYPOTHETICAL 47.6 KD PROTEIN IN EPR-GALK INTERGENIC REGION [Bacillus subtilis]
621	Bt1Gc911	Bt1G805	1405-1	g2108273	269	399	4.00E-37	43	52	(X99724) EA1 [Bacillus anthracis]
622	Bt1Gc912	Bt1G806	1-760	g2632005	958	593	1.10E-57	77	77	(X92868) NAD alcohol dehydrogenase [Bacillus subtilis]
622	Bt1Gc912	Bt1G807	895-1625	g2635735	676	435	6.10E-41	54	91	(AJ002571) YkaB [Bacillus subtilis]
623	Bt1Gc913	Bt1G808	1-1676	g282243	47	192	2.40E-15	27	16	(Z99120) yunE [Bacillus subtilis]
624	Bt1Gc910	Bt1G809	822-1	g2688027	183	189	1.70E-13	22	27	collagen adhesin - Staphylococcus aureus [] (AE001125) acriflavin resistance protein (acrB) [Borrelia burgdorferi]
624	Bt1Gc910	Bt1G810	825-503	g128506	132	157	3.40E-11	29	39	NODULATION PROTEIN NOLG PRECURSOR

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
625	Bt1Gc915	Bt1G811	1-762	g3114663	428	312	6.60E-28	37	90	[Rhizobium meliloti] (AF061267) inner membrane component HtxC [Pseudomonas stutzeri]
625	Bt1Gc915	Bt1G812	1-720	g3127079	412	329	1.00E-29	37	86	(AF061070) PxC [Pseudomonas stutzeri]
626	Bt1Gc919	Bt1G813	177-1	g134774	213	168	5.00E-12	69	16	STAGE V SPORULATION PROTEIN E [Bacillus subtilis]
626	Bt1Gc919	Bt1G814	1611-262	g417327	1279	1300	1.30E-132	55	100	UDP-N-ACETYLMURAMOYLALANINE--D-GLUTAMATE LIGASE (UDP-N-ACETYLMURAMOYL-L-ALANYL-D-GLUTAMATE SYNTHETASE) (D-GLUTAMIC ACID ADDING ENZYME) [Bacillus subtilis]
626	Bt1Gc919	Bt1G815	1942-1615	g417313	475	394	1.30E-36	85	34	PHOSPHO-N-ACETYLMURAMOYL-PENTAPEPTIDE-TRANSFERASE (UDP-MURNAC-PENTAPEPTIDE PHOSPHOTRANSFERASE) [Bacillus subtilis]
627	Bt1Gc917	Bt1G816	1-243	g4584149	274	280	1.60E-24	75	26	(AJ010111) cytochrome aa3 controlling protein [Bacillus cereus]
627	Bt1Gc917	Bt1G817	1801-334	g4584148	2453	2175	2.50E-225	97	50	(AJ010111) pyruvate carboxylase [Bacillus cereus]
628	Bt1Gc920	Bt1G818	1444-866	g2313950	119	136	5.70E-18	25	100	(AE000593) osmoprotection protein (proV) [Helicobacter pylori 26695]
629	Bt1Gc921	Bt1G819	1-1657	g2621914	217	436	1.20E-40	29	52	(AE000859) sensory transduction histidine kinase [Methanobacterium thermoautotrophicum]
630	Bt1Gc922	Bt1G820	684-1361	g1684650	380	415	8.00E-39	39	47	(Z82987) unknown [Bacillus subtilis]
631	Bt1Gc923	Bt1G821	554-1	g2633731	550	524	2.30E-50	53	78	(Z99111) similar to hypothetical proteins [Bacillus subtilis]
631	Bt1Gc923	Bt1G822	1813-554	g2633730	1212	1288	2.50E-131	60	100	(Z99111) similar to ribulose-bisphosphate carboxylase [Bacillus subtilis]
632	Bt1Gc924	Bt1G823	1-1227	g417328	1188	1154	3.90E-117	57	83	UDP-N-ACETYLMURAMOYLALANYL-D-GLUTAMATE--2,6-DIAMINOPIMELATE LIGASE (UDP-N-ACETYLMURAMYL-TRIPETIDE SYNTHETASE) [Bacillus subtilis]
632	Bt1Gc924	Bt1G824	1246-1901	g417313	764	639	1.50E-62	68	67	PHOSPHO-N-ACETYLMURAMOYL-PENTAPEPTIDE-TRANSFERASE (UDP-MURNAC-PENTAPEPTIDE PHOSPHOTRANSFERASE) [Bacillus subtilis]
633	Bt1Gc928	Bt1G825	469-115	g585905	337	365	1.60E-33	59	95	RIBONUCLEASE P PROTEIN COMPONENT (PROTEIN C5) (RNASE P) [Bacillus subtilis]
633	Bt1Gc928	Bt1G826	680-549	g132902	204	204	1.80E-16	91	98	50S RIBOSOMAL PROTEIN L34 [Bacillus stearothermophilus]
634	Bt1Gc926	Bt1G827	1523-1093	g2224769	394	318	1.50E-28	55	69	(Z97025) product similar to B. subtilis YhcN hypothetical



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
635	Bt1Gc927	Bt1G828	1521-955	g3322295	169	101	0.0018	27	54	21.0 KD lipoprotein [Bacillus subtilis] (AE001189) regulatory protein (pfoS/R) [Treponema pallidum]
636	Bt1Gc929	Bt1G829	31-402	g2632455	492	435	6.10E-41	77	99	(Z99104) ybcI [Bacillus subtilis]
636	Bt1Gc929	Bt1G830	1001-1548	g1171678	630	526	1.40E-50	69	36	NADH DEHYDROGENASE SUBUNIT 5 (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5) [Bacillus subtilis]
637	Bt1Gc931	Bt1G831	1033-1274	g1176282	156	198	1.80E-15	46	23	HYPOTHETICAL OXIDOREDUCTASE IN GNTR-GGT INTERGENIC REGION [Escherichia coli]
638	Bt1Gc930	Bt1G832	1142-876	g1161061	196	216	9.80E-18	49	99	(L43135) dioxygenase [Methylobacterium extorquens]
638	Bt1Gc930	Bt1G833	1142-210	g2632008	753	789	1.90E-78	49	100	(AJ002571) YkcA [Bacillus subtilis]
639	Bt1Gc934	Bt1G834	418-1	g2896708	240	282	1.00E-24	40	40	(AL021897) echA9 [Mycobacterium tuberculosis]
639	Bt1Gc934	Bt1G835	1441-1720	g548899	187	198	1.30E-15	39	29	EXONUCLEASE SBCD HOMOLOG []
640	Bt1Gc932	Bt1G836	1833-874	g2619033	1064	1026	1.40E-103	66	37	(AF027868) PEP synthase [Bacillus subtilis]
641	Bt1Gc916	Bt1G837	Jan-09	g1709698	2674	2655	3.40E-276	83	90	POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE (POLYNUCLEOTIDE PHOSPHORYLASE) (PNPASE) (VEGETATIVE PROTEIN 15) (VEG15) [Bacillus subtilis]
642	Bt1Gc935	Bt1G838	427-750	g2495742	85	114	1.10E-06	26	50	HYPOTHETICAL PROTEIN MJ0014 [Methanococcus jannaschii]
642	Bt1Gc935	Bt1G839	710-1774	g1002992	320	441	1.40E-41	35	90	(U34772) ORF375 [Dichelobacter nodosus]
642	Bt1Gc935	Bt1G840	704-1774	g3426013	288	362	3.30E-33	28	85	(AB016803) transposase [Deinococcus radiodurans]
642	Bt1Gc935	Bt1G841	1487-1774	g1149666	269	271	1.50E-23	57	78	(X86498) IS1136 DNA [Clostridium perfringens]
643	Bt1Gc933	Bt1G842	1-1193	g2633857	1676	1578	4.60E-162	79	35	(Z99111) pyruvate carboxylase [Bacillus subtilis]
644	Bt1Gc936	Bt1G843	1-169	g1789981	110	135	1.10E-08	44	20	(AE000433) IS150 putative transposase [Escherichia coli]
645	Bt1Gc939	Bt1G844	389-1	g1709639	441	458	2.20E-43	63	32	PEPTIDASE T (AMINOTRIPEPTIDASE) (TRIPEPTIDASE) [Bacillus subtilis]
646	Bt1Gc937	Bt1G845	738-1	g1805447	231	279	2.10E-24	28	65	(D50453) response-regulator aspartate phosphatase C [Bacillus subtilis]
647	Bt1Gc938	Bt1G846	1282-1026	g2114014	84	108	3.20E-05	32	16	(Z95558) hypothetical protein Rv0552 [Mycobacterium tuberculosis]
648	Bt1Gc941	Bt1G847	701-1	g732327	720	605	5.90E-59	60	75	PROBABLE 1,4-DIHYDROXY-2-NAPHTHOATE OCTAPRENYLTRANSFERASE (DHNA-OCTAPRENYLTRANSFERASE) [Bacillus subtilis]
649	Bt1Gc914	Bt1G848	1587-1	g2116756	1609	1628	2.30E-167	59	84	(D86418) YfmR [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
649	Bt1Gc914	Bt1G849	2231-1580	g1730910	353	392	2.20E-36	41	100	HYPOTHETICAL 23.1 KD PROTEIN IN BSAA-ILVD INTERGENIC REGION [Bacillus subtilis]
650	Bt1Gc942	Bt1G850	201-889	g2833393	560	513	3.30E-49	47	100	SAPB PROTEIN [Bacillus subtilis]
651	Bt1Gc943	Bt1G851	1-465	g1934657	153	201	8.20E-16	43	23	(U93876) hypothetical protein YrdP [Bacillus subtilis]
652	Bt1Gc947	Bt1G852	235-1010	g1763710	263	293	6.80E-26	31	100	(Z83337) highly similar to phosphotransferase system regulator [Bacillus subtilis]
652	Bt1Gc947	Bt1G853	552-1549	g2621392	75	175	2.20E-13	28	99	(AE000818) conserved protein [Methanobacterium thermoautotrophicum]
653	Bt1Gc945	Bt1G854	456-1	g1934810	358	361	4.20E-33	49	64	(Z93937) unknown [Bacillus subtilis]
653	Bt1Gc945	Bt1G855	1210-453	g1934809	554	590	2.30E-57	45	47	(Z93937) unknown [Bacillus subtilis]
654	Bt1Gc944	Bt1G856	1552-1	g3915461	1425	1419	3.30E-145	50	89	HYPOTHETICAL 66.6 KD PROTEIN IN PURD-SAPB INTERGENIC REGION [Bacillus subtilis]
655	Bt1Gc949	Bt1G857	2158-1397	g3059028	311	301	2.80E-26	32	50	(Y12537) proP [Corynebacterium glutamicum]
656	Bt1Gc948	Bt1G858	432-1	g2811132	402	423	1.10E-39	56	26	(AF013216) Mls [Myxococcus xanthus]
656	Bt1Gc948	Bt1G859	1281-1	g4894306	455	539	5.80E-52	49	34	(AF065404) pXO1-90 [Bacillus anthracis]
657	Bt1Gc951	Bt1G860	1797-1221	g4914333	418	398	5.10E-37	46	80	(AC005489) F14N23.19 [Arabidopsis thaliana]
658	Bt1Gc946	Bt1G861	880-34	g1171920	801	616	4.00E-60	59	100	GLYCINE BETAIN TRANSPORT SYSTEM PERMEASE PROTEIN OPUAB [Bacillus subtilis]
658	Bt1Gc946	Bt1G862	1891-839	g1805370	1037	968	2.00E-97	60	84	(D50453) glycine betain/L-proline transport ATPase [Bacillus subtilis]
659	Bt1Gc953	Bt1G863	439-1209	g115022	1294	1179	8.80E-120	98	100	BETA-LACTAMASE PRECURSOR, TYPE II (PENICILLINASE) (CEPHALOSPORINASE) [Bacillus cereus]
660	Bt1Gc955	Bt1G864	1710-1785	g2497382	109	125	3.40E-07	96	6	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
661	Bt1Gc954	Bt1G865	1-140	g2634845	116	122	9.00E-08	48	60	(Z99116) yqzF [Bacillus subtilis]
661	Bt1Gc954	Bt1G866	918-208	g1731058	495	459	1.70E-43	43	100	HYPOTHETICAL 27.0 KD PROTEIN IN SPO0A-MMGA INTERGENIC REGION [Bacillus subtilis]
661	Bt1Gc954	Bt1G867	842-1098	g586811	135	162	5.20E-12	38	97	HYPOTHETICAL 10.2 KD PROTEIN IN DNAC-RPLI INTERGENIC REGION [Bacillus subtilis]
662	Bt1Gc957	Bt1G868	Jan-16	g1731087	171	260	2.10E-22	30	59	HYPOTHETICAL 39.0 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
663	Bt1Gc956	Bt1G869	66-1889	g1881242	2068	1683	3.40E-173	66	100	(AB001488) FUNCTION UNKNOWN, WEAK SIMILARITY TO YEZF_ECOLI. [Bacillus subtilis]
664	Bt1Gc958	Bt1G870	1386-708	g2621255	150	191	4.40E-15	24	100	(AE000808) conserved protein [Methanobacterium thermoautotrophicum]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
665	Bt1Gc961	Bt1G871	1-111	g2633504	129	129	1.60E-08	65	27	thermoautotrophicum]
665	Bt1Gc961	Bt1G872	428-906	g2982976	86	140	1.10E-09	22	99	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
666	Bt1Gc960	Bt1G873	987-486	g135583	488	517	1.20E-49	56	71	hypothetical protein [Aquifex aeolicus]
667	Bt1Gc963	Bt1G874	88-1361	g417332	103	131	1.00E-08	38	61	TRANSCRIPTIONAL ACTIVATOR TENA [Bacillus subtilis]
667	Bt1Gc963	Bt1G874	88-1361	g417332	103	131	1.00E-08	38	61	MUTT-LIKE PROTEIN (ORF154) [Streptomyces ambifaciens]
667	Bt1Gc963	Bt1G875	1136-552	g1770039	164	197	1.00E-15	23	99	(Z75208) hypothetical protein [Bacillus subtilis]
668	Bt1Gc964	Bt1G876	340-939	g732345	625	626	3.50E-61	59	100	HYPOTHETICAL 22.6 KD PROTEIN IN EPR-GALK INTERGENIC REGION [Bacillus subtilis]
669	Bt1Gc965	Bt1G877	1-836	g3183483	1327	1190	6.00E-121	95	89	HYPOTHETICAL PROTEIN (ORF3) [Bacillus cereus]
669	Bt1Gc965	Bt1G878	1246-1917	g1075694	385	419	3.00E-39	38	42	pheromone CAD1 binding protein precursor - Enterococcus faecalis plasmid pAD1 [Plasmid pAD1]
670	Bt1Gc969	Bt1G879	880-1	g418534	625	676	1.80E-66	54	88	HYPOTHETICAL 32.5 KD PROTEIN IN PEPE-LYSC INTERGENIC REGION [Escherichia coli]
671	Bt1Gc967	Bt1G880	105-1084	g2626830	1161	1211	3.60E-123	65	58	(D83967) TreC [Bacillus subtilis]
672	Bt1Gc966	Bt1G881	458-1	g3688811	604	568	4.90E-55	80	42	(AF084104) maltose transporter ATP-binding protein [Bacillus firmus]
672	Bt1Gc966	Bt1G882	721-2268	g129005	2733	2683	3.70E-279	97	92	OLIGO-1,6-GLUCOSIDASE (SUCRASE-ISOMALTASE) (LIMIT DEXTRINASE) (ISOMALTASE) (DEXTRIN 6-ALPHA-D-GLUCANOHYDROLASE) [Bacillus cereus]
673	Bt1Gc970	Bt1G883	1-398	g2126635	80	145	3.30E-10	30	97	gyrb protein - Streptococcus pneumoniae (fragment) [Streptococcus pneumoniae]
673	Bt1Gc970	Bt1G884	1313-1076	g2634267	296	307	2.20E-27	79	81	(Z99114) similar to transcriptional regulator [Bacillus subtilis]
673	Bt1Gc970	Bt1G885	1806-1327	g2634266	618	571	2.40E-55	74	99	(Z99114) yoaS [Bacillus subtilis]
674	Bt1Gc972	Bt1G886	1-426	g1730951	423	423	1.10E-39	54	84	HYPOTHETICAL PHOSPHOTRANSFERASE ENZYME II, A COMPONENT IN PONA-COTD INTERGENIC REGION (F143) [Bacillus subtilis]
674	Bt1Gc972	Bt1G887	675-1115	g732386	279	306	2.90E-27	43	99	PUTATIVE LOW MOLECULAR WEIGHT PROTEIN-TYROSINE-PHOSPHATASE [Bacillus subtilis]
674	Bt1Gc972	Bt1G888	1210-1347	g732387	180	180	6.40E-14	65	30	HYPOTHETICAL LACA/RPIB FAMILY PROTEIN IN SPOIIR-GLYC INTERGENIC REGION [Bacillus subtilis]
675	Bt1Gc971	Bt1G889	41-1336	g421457	1600	1524	2.40E-156	71	100	secY protein - Bacillus licheniformis []
676	Bt1Gc973	Bt1G890	885-268	g2632848	564	568	4.90E-55	54	100	(Z99106) similar to hypothetical proteins [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
677	Bt1Gc975	Bt1G891	1013-1	g1175322	891	684	2.50E-67	53	74	HYPOTHETICAL PROTEIN HI0883 [Haemophilus influenzae Rd]
678	Bt1Gc976	Bt1G892	1-706	g2619000	814	831	6.60E-83	62	89	(AF027868) YoaT [Bacillus subtilis]
678	Bt1Gc976	Bt1G893	1800-813	g2522022	546	559	4.40E-54	39	100	(AB007638) peroxidase [Bacillus subtilis]
679	Bt1Gc974	Bt1G894	542-1	g730600	668	595	6.80E-58	66	38	AMINO-ACID PERMEASE ROCC [Bacillus subtilis]
679	Bt1Gc974	Bt1G895	1394-747	g2635191	733	632	8.10E-62	68	100	(Z99117) glutamine ABC transporter (integral membrane protein) [Bacillus subtilis]
679	Bt1Gc974	Bt1G896	1388-738	g2635190	396	356	1.40E-32	41	100	(Z99117) glutamine ABC transporter (integral membrane protein) [Bacillus subtilis]
680	Bt1Gc979	Bt1G897	841-1566	g1881262	633	638	1.90E-62	48	63	(AB001488) SIMILAR TO ACYL-COA DEHYDROGENASE. [Bacillus subtilis]
681	Bt1Gc980	Bt1G898	266-1	g420808	337	279	2.10E-24	70	88	hypothetical 12K protein - Thermus aquaticus [Thermus aquaticus]
682	Bt1Gc981	Bt1G899	53-904	g2635714	1008	923	1.20E-92	66	100	(Z99120) similar to diaminopimelate epimerase [Bacillus subtilis]
682	Bt1Gc981	Bt1G900	952-1159	g2635713	166	98	3.10E-05	44	57	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
683	Bt1Gc982	Bt1G901	709-116	g2497586	299	391	2.80E-36	41	71	D-ALANYL-D-ALANINE CARBOXYPEPTIDASE (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) [Enterococcus faecalis]
684	Bt1Gc984	Bt1G902	1-1415	g1169000	1250	1318	1.60E-134	56	38	MICROBIAL COLLAGENASE PRECURSOR (120 KD COLLAGENASE) [Clostridium perfringens]
685	Bt1Gc983	Bt1G903	360-1	g4105715	189	233	1.60E-19	44	47	(AF050754) glucose 6-phosphate isomerase [Giardia intestinalis]
685	Bt1Gc983	Bt1G904	1510-353	g3122425	711	747	5.30E-74	42	100	N-ACETYLGLUCOSAMINE-6-PHOSPHATE DEACETYLASE (GLCNAC 6-P DEACETYLASE) [Bacillus subtilis]
686	Bt1Gc985	Bt1G905	812-1	g131027	398	458	2.20E-43	42	92	PROTEIN EXPORT PROTEIN PRSA PRECURSOR [Bacillus subtilis]
687	Bt1Gc987	Bt1G906	324-1	g401655	246	249	3.10E-21	49	57	MAF PROTEIN [Bacillus subtilis]
688	Bt1Gc988	Bt1G907	174-1223	g3915557	908	756	5.90E-75	55	100	HYPOTHETICAL LIPOPROTEIN YUFN PRECURSOR [Bacillus subtilis]
689	Bt1Gc986	Bt1G908	1862-135	g2226222	346	430	2.10E-40	38	87	(Y14082) hypothetical protein [Bacillus subtilis]
689	Bt1Gc986	Bt1G909	1862-1034	g1575578	303	328	1.30E-29	31	69	(U67196) histidine protein kinase [Thermotoga maritima]
690	Bt1Gc989	Bt1G910	1-965	g2245638	1399	1265	6.80E-129	82	94	(AF006075) TPP-dependent acetoin dehydrogenase, E1 beta-subunit [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
690	Bt1Gc989	Bt1G911	1002-1545	g2633132	376	384	1.50E-35	53	46	(Z99108) acetoin dehydrogenase E2 component (dihydrolipoamide acetyltransferase) [Bacillus subtilis]
691	Bt1Gc992	Bt1G912	1-1128	g1694898	597	635	1.70E-61	36	37	(Y09450) transposase [Pseudomonas putida]
691	Bt1Gc992	Bt1G913	1-1128	g79972	194	299	2.70E-25	26	34	transposase tnpA - Enterococcus faecalis plasmid pAD2 transposon Tn917 [Transposon Tn917]
692	Bt1Gc990	Bt1G914	1231-1602	g3915567	496	458	2.20E-43	73	99	HYPOTHETICAL 14.8 KD PROTEIN IN HOM-MRGA INTERGENIC REGION [Bacillus subtilis]
692	Bt1Gc990	Bt1G915	1984-2267	g2145816	117	144	4.20E-10	28	99	bacitracin resistance protein homolog bacA - Mycobacterium leprae [Mycobacterium leprae]
693	Bt1Gc993	Bt1G916	1-1020	g2415733	210	260	2.10E-22	24	94	(AB000617) RapJ [Bacillus subtilis]
694	Bt1Gc991	Bt1G917	1228-211	g1665720	1623	1644	4.70E-169	95	93	(D17312) diarrheal toxin [Bacillus cereus]
695	Bt1Gc995	Bt1G918	52-945	g118407	452	511	5.40E-49	37	100	D-ALANINE--D-ALANINE LIGASE B (D-ALANYLALANINE SYNTHETASE) [Escherichia coli]
695	Bt1Gc995	Bt1G919	1278-1509	g135583	195	221	2.90E-18	56	33	TRANSCRIPTIONAL ACTIVATOR TENA [Bacillus subtilis]
696	Bt1Gc997	Bt1G920	1-214	g2635183	193	162	2.20E-11	58	20	(Z99117) similar to folate metabolism [Bacillus subtilis]
696	Bt1Gc997	Bt1G921	406-1053	g2635182	559	580	2.60E-56	50	100	(Z99117) similar to caffeoyl-CoA O-methyltransferase [Bacillus subtilis]
697	Bt1Gc998	Bt1G922	992-1	g1945051	1399	904	1.20E-104	84	86	(U63928) L1 protein [Bacillus cereus]
697	Bt1Gc998	Bt1G923	1943-1032	g1945050	1074	902	2.00E-90	72	68	(U63928) L2 protein [Bacillus cereus]
698	Bt1Gc1001	Bt1G924	723-151	g2983883	170	225	1.10E-18	30	100	(AE000743) O-methyltransferase [Aquifex aeolicus]
699	Bt1Gc1000	Bt1G925	1-650	g3183575	541	509	8.80E-49	47	89	HYPOTHETICAL 28.1 KD PROTEIN IN PHOD-PCP INTERGENIC REGION (ORF19) [Bacillus subtilis]
699	Bt1Gc1000	Bt1G926	995-1243	g1730936	127	127	2.60E-08	31	99	HYPOTHETICAL 10.2 KD PROTEIN IN ILVA 3' REGION [Bacillus subtilis]
700	Bt1Gc1002	Bt1G927	1-1405	g3329623	52	140	4.80E-09	25	43	(AF078790) No definition line found [Caenorhabditis elegans]
701	Bt1Gc996	Bt1G928	1077-415	g2822199	490	534	2.00E-51	45	100	(U96166) ATP-binding cassette protein [Streptococcus crista]
702	Bt1Gc999	Bt1G929	1-514	g1842440	406	443	8.70E-42	53	41	(U87792) ClnA [Bacillus subtilis]
702	Bt1Gc999	Bt1G930	662-1705	g72992	1470	1401	2.60E-143	83	100	recE protein - Bacillus subtilis []
703	Bt1Gc1003	Bt1G931	739-1	g1731038	204	183	3.10E-14	52	38	HYPOTHETICAL 28.3 KD PROTEIN IN COMGG-SINR INTERGENIC REGION [Bacillus subtilis]
703	Bt1Gc1003	Bt1G932	1357-787	g1731037	414	366	1.20E-33	44	99	SIGNAL PEPTIDASE I W (SPASE I) (LEADER PEPTIDASE I) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
704	Bt1Gc1005	Bt1G933	1772-225	g2266421	1738	1787	3.30E-184	65	100	(Y13917) yngK [Bacillus subtilis]
705	Bt1Gc1006	Bt1G934	772-1	g416901	1007	866	1.30E-86	77	74	ASPARTATE-SEMIALDEHYDE DEHYDROGENASE (ASA DEHYDROGENASE) [Bacillus subtilis]
705	Bt1Gc1006	Bt1G935	1531-932	g417831	785	792	9.00E-79	74	100	DIPICOLINATE SYNTHASE, B CHAIN [Bacillus subtilis]
706	Bt1Gc1007	Bt1G936	1691-1151	g1731048	607	609	2.20E-59	65	51	PUTATIVE PEPTIDASE IN GCVT-SPOIIIAA INTERGENIC REGION [Bacillus subtilis]
707	Bt1Gc1008	Bt1G937	690-1	g2116761	613	589	2.90E-57	53	57	(D86418) YfnC [Bacillus subtilis]
707	Bt1Gc1008	Bt1G938	1272-893	g3868866	82	113	1.10E-05	23	22	(AB014075) GTPase [Clostridium histolyticum]
708	Bt1Gc1004	Bt1G939	167-1478	g2688567	642	614	6.60E-60	37	100	(AE001165) Na <sup>+</sup> /H <sup>+</sup> antiporter (nhaC-1) [Borrelia burgdorferi]
709	Bt1Gc1009	Bt1G940	1435-323	g3915568	1359	1369	6.50E-140	71	100	HYPOTHETICAL 39.4 KD OXIDOREDUCTASE IN HOM-MRGA INTERGENIC REGION [Bacillus subtilis]
710	Bt1Gc1010	Bt1G941	1821-1563	g2632030	224	234	1.20E-19	53	82	(AJ002571) YkkD [Bacillus subtilis]
711	Bt1Gc1011	Bt1G942	1245-1	g3876615	805	801	1.00E-79	43	48	(Z74031) Similarity to Yeast D-lactate dehydrogenase (SW:DL1_YEAST); cDNA EST EMBL:C12235 comes from this gene; cDNA EST EMBL:C10532 comes from this gene; cDNA EST EMBL:C10979 comes f... []
712	Bt1Gc1012	Bt1G943	1-646	g2274944	529	479	1.30E-45	46	54	(AJ00346) NapC protein [Enterococcus hirae]
713	Bt1Gc1013	Bt1G944	347-1	g225559	560	572	1.90E-55	98	24	ORF IS231C [Bacillus thuringiensis]
713	Bt1Gc1013	Bt1G945	1794-1349	g2497392	754	753	1.20E-74	98	59	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
714	Bt1Gc1014	Bt1G946	1832-1	g2635858	278	358	7.60E-32	38	26	(Z99121) yvgS [Bacillus subtilis]
714	Bt1Gc1014	Bt1G947	1770-1535	g732400	76	137	6.30E-09	34	29	HYPOTHETICAL 31.7 KD PROTEIN IN KINC-ADEC INTERGENIC REGION (ORF3) [Bacillus subtilis]
714	Bt1Gc1014	Bt1G948	2462-2141	g2619051	236	272	7.60E-23	52	18	(AF027868) RecQ homolog [Bacillus subtilis]
715	Bt1Gc1015	Bt1G949	135-548	g3676415	89	131	1.00E-08	22	99	(AF051917) unknown [Staphylococcus aureus]
716	Bt1Gc1017	Bt1G950	167-2407	g3130015	142	227	4.10E-18	26	49	(AL023517) putative iron-sulfur binding oxidoreductase [Streptomyces coelicolor]
717	Bt1Gc1016	Bt1G951	1-1596	g1171759	501	774	7.30E-77	35	93	BACILLOLYSIN PRECURSOR (NEUTRAL PROTEASE) [Bacillus brevis]
717	Bt1Gc1016	Bt1G952	1598-1864	g2696017	166	204	3.10E-15	44	11	(AB001874) chitinase B [Clostridium paraputrificum]
718	Bt1Gc1018	Bt1G953	23-711	g4584144	949	717	8.00E-71	76	100	(AJ010140) DNA alkylation repair enzyme [Bacillus cereus]
719	Bt1Gc1019	Bt1G954	1-173	g2340003	156	130	1.30E-08	58	62	(Z98682) YlbG protein [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
719	Bt1Gc1019	Bt1G955	1110-1601	g2340004	517	518	9.80E-50	61	99	(Z98682) YlbH protein [Bacillus subtilis]
719	Bt1Gc1019	Bt1G956	1673-1921	g2340005	310	310	1.10E-27	69	51	(Z98682) YlbI protein [Bacillus subtilis]
720	Bt1Gc1021	Bt1G957	144-839	g2633027	595	609	2.20E-59	50	100	(Z99107) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
721	Bt1Gc1023	Bt1G958	1317-824	g2618995	426	402	1.90E-37	50	99	(AF027868) putative alanine acetyl transferase [Bacillus subtilis]
721	Bt1Gc1023	Bt1G959	1580-1338	g1770024	245	263	1.00E-22	56	99	(Z75208) hypothetical protein [Bacillus subtilis]
722	Bt1Gc1025	Bt1G960	1-140	g2632987	172	187	7.10E-14	71	10	(Z99107) alternate gene name; yerS; similar to RNA methyltransferase [Bacillus subtilis]
722	Bt1Gc1025	Bt1G961	1357-187	g4490997	364	442	1.10E-41	29	100	(AL035707) putative integrase [Streptomyces coelicolor]
723	Bt1Gc1024	Bt1G962	445-1	g1399821	231	238	4.60E-20	38	54	(U59229) PhoC [Rhizobium meliloti]
723	Bt1Gc1024	Bt1G963	2269-532	g1399822	62	150	2.80E-10	24	63	(U59229) PhoD precursor [Rhizobium meliloti]
723	Bt1Gc1024	Bt1G964	1014-2269	g2626826	199	306	9.00E-26	36	13	(D83967) YfkN [Bacillus subtilis]
724	Bt1Gc1027	Bt1G965	508-242	g2507327	405	405	9.20E-38	84	99	30S RIBOSOMAL PROTEIN S15 (BS18) [Bacillus subtilis]
724	Bt1Gc1027	Bt1G966	1330-636	g1592690	664	691	4.50E-68	54	73	(Z80835) riboflavin kinase [Bacillus subtilis]
725	Bt1Gc1031	Bt1G967	1-635	g2633840	761	760	2.20E-75	67	69	(Z99111) ykD [Bacillus subtilis]
725	Bt1Gc1031	Bt1G968	339-1971	g4678913	146	370	4.70E-34	29	78	(AL049707) hypothetical protein [Streptomyces coelicolor]
726	Bt1Gc1029	Bt1G969	1-148	g2500744	118	93	0.0003	48	22	REGULATORY PROTEIN VANRB [Enterococcus faecalis]
726	Bt1Gc1029	Bt1G970	159-1589	g2500765	396	404	1.20E-37	28	100	SENSOR PROTEIN VANSB (VANCOMYCIN B-TYPE RESISTANCE PROTEIN VANSB) (VANCOMYCIN HISTIDINE PROTEIN KINASE) [Enterococcus faecalis]
727	Bt1Gc1032	Bt1G971	1457-212	g1175719	863	943	9.00E-95	44	62	PENICILLIN-BINDING PROTEIN 3 (PBP 3) [Bacillus subtilis]
728	Bt1Gc1034	Bt1G972	1-398	g118797	242	268	3.00E-23	46	35	DNA POLYMERASE III, BETA CHAIN [Bacillus subtilis]
729	Bt1Gc1039	Bt1G973	252-1559	g131733	1903	1909	3.90E-197	83	81	CTP SYNTHASE (UTP--AMMONIA LIGASE) (CTP SYNTHETASE) [Bacillus subtilis]
730	Bt1Gc1036	Bt1G974	211-1210	g2633169	976	776	4.50E-77	58	100	(Z99108) similar to iron(III) dicitrate transport permease [Bacillus subtilis]
730	Bt1Gc1036	Bt1G975	1236-2074	g2633170	683	594	8.60E-58	52	81	(Z99108) similar to iron(III) dicitrate transport permease [Bacillus subtilis]
731	Bt1Gc1028	Bt1G976	89-1408	g2239289	1504	812	7.60E-105	66	100	(U51115) unknown protein [Bacillus subtilis]
732	Bt1Gc1038	Bt1G977	2107-2476	g1644192	598	614	6.60E-60	98	47	(D63645) spore cortex-lytic enzyme prepeptide [Bacillus cereus]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
733	Bt1Gc1033	Bt1G978	1-607	g586024	868	875	1.40E-87	79	43	STAGE V SPOULATION PROTEIN R [Bacillus subtilis]
734	Bt1Gc1037	Bt1G979	663-318	g2633181	138	174	2.80E-13	30	99	(Z99108) yfH [Bacillus subtilis]
734	Bt1Gc1037	Bt1G980	1464-787	g1709315	262	337	1.50E-30	36	100	CHITOLIGOSACCHARIDE DEACETYLASE (NODULATION PROTEIN B) [Rhizobium galegae]
735	Bt1Gc1035	Bt1G981	339-1292	g2649231	131	286	3.80E-25	25	100	(AE001011) conserved hypothetical protein [Archaeoglobus fulgidus]
736	Bt1Gc1040	Bt1G982	1-1300	g1730990	1868	1877	9.50E-194	81	96	HYPOTHETICAL 51.7 KD PROTEIN IN DNAJ-RPSU INTERGENIC REGION [Bacillus subtilis]
737	Bt1Gc1041	Bt1G983	1499-9	g3249555	1582	1502	5.20E-154	61	100	(AF010184) coenzyme A transferase PscA [Pseudomonas aeruginosa]
737	Bt1Gc1041	Bt1G984	2094-1636	g1770002	141	180	6.40E-14	33	54	(Z75208) autolysin response regulator [Bacillus subtilis]
738	Bt1Gc1043	Bt1G985	421-1	g266568	599	636	3.10E-62	88	41	ROD SHAPE-DETERMINING PROTEIN MREB [Bacillus subtilis]
738	Bt1Gc1043	Bt1G986	1274-582	g400964	774	756	5.90E-75	65	100	DNA REPAIR PROTEIN RADC HOMOLOG (ORFB) [Bacillus subtilis]
738	Bt1Gc1043	Bt1G987	1635-1315	g401655	304	315	3.20E-28	54	56	MAF PROTEIN [Bacillus subtilis]
739	Bt1Gc1042	Bt1G988	1-709	g267022	879	639	1.50E-62	70	88	STAGE III SPOULATION PROTEIN J PRECURSOR [Bacillus subtilis]
739	Bt1Gc1042	Bt1G989	709-1153	g400078	305	333	3.90E-30	42	71	JAG PROTEIN (SPOIII ASSOCIATED PROTEIN) [Bacillus subtilis]
740	Bt1Gc1045	Bt1G990	1-1154	g1168447	491	482	6.40E-46	33	91	XAA-PRO AMINOPEPTIDASE (X-PRO AMINOPEPTIDASE) (AMINOPEPTIDASE P II) (APP-II) (AMINOACYLPROLINE AMINOPEPTIDASE) [Haemophilus influenzae Rd]
740	Bt1Gc1045	Bt1G991	1935-1297	g2145390	328	318	1.50E-28	33	100	(Y09476) YisX [Bacillus subtilis]
741	Bt1Gc1047	Bt1G992	1509-611	g461503	798	813	5.40E-81	53	100	ALS OPERON REGULATORY PROTEIN [Bacillus subtilis]
742	Bt1Gc1046	Bt1G993	2820-348	g746118	670	764	8.30E-76	34	100	DNA photolyase [Bacillus firmus]
742	Bt1Gc1046	Bt1G994	1772-1	g2341035	753	857	1.20E-85	37	80	(AC000104) Match to Arabidopsis photolysase (PHI1) gene (gb X99061) and cryptochrome 2 apoprotein (CRY2) (gb U43397). ESTs gb W43661 and gb Z25638 come from this gene. [Arabidopsis thaliana]
742	Bt1Gc1046	Bt1G995	2676-1848	g1001807	277	326	2.20E-29	31	100	(D64005) hypothetical protein [Synechocystis sp.]
743	Bt1Gc1044	Bt1G996	1-1022	g1172061	443	502	4.80E-48	34	70	AMINOACYL-HISTIDINE DIPEPTIDASE (XAA-HIS DIPEPTIDASE) (X-HIS DIPEPTIDASE) (BETA-



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat. nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
744	Bt1Gc1048	Bt1G997	1533-1	g124464	2075	1939	2.60E-200	74	74	ALANYL-HISTIDINE DIPEPTIDASE (CARNOSINASE) (PEPTIDASE D) [Haemophilus influenzae Rd]
745	Bt1Gc1049	Bt1G998	102-1	g2340003	132	132	7.80E-09	68	37	IMMUNE INHIBITOR A PRECURSOR [Bacillus thuringiensis]
745	Bt1Gc1049	Bt1G999	678-235	g2340002	462	397	6.50E-37	59	99	(Z98682) YlbG protein [Bacillus subtilis]
745	Bt1Gc1049	Bt1G1000	1368-790	g2226254	321	327	1.70E-29	38	99	(Y14084) hypothetical protein [Bacillus subtilis]
745	Bt1Gc1049	Bt1G1001	1690-1454	g2340001	204	204	1.80E-16	44	99	(Z98682) YlbE protein [Bacillus subtilis]
746	Bt1Gc1053	Bt1G1002	25-594	g127481	790	794	5.50E-79	79	99	GTP CYCLOHYDROLASE I (GTP-CH-I) [Bacillus subtilis]
746	Bt1Gc1053	Bt1G1003	778-1562	g399774	370	353	3.00E-32	42	98	PROBABLE HEPTAPRENYL DIPHOSPHATE SYNTHASE COMPONENT I (HEPPP SYNTHASE) (SPORE GERMINATION PROTEIN C1) [Bacillus subtilis]
747	Bt1Gc1050	Bt1G1004	950-1	g1805408	1061	996	2.20E-100	64	80	(D50453) homologues to nitrile hydratase region 3'-hypothetical protein P47K of P. chlororaphis [Bacillus subtilis]
747	Bt1Gc1050	Bt1G1005	1239-1560	g1805397	293	305	3.60E-27	52	36	(D50453) homologue of unidentified protein of E. coli [Bacillus subtilis]
748	Bt1Gc1051	Bt1G1006	265-11	g2634060	254	262	1.30E-22	54	99	(Z99112) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
748	Bt1Gc1051	Bt1G1007	1054-344	g2634059	613	586	6.10E-57	52	100	(Z99112) similar to 3-oxoacyl-acyl-carrier protein reductase [Bacillus subtilis]
748	Bt1Gc1051	Bt1G1008	1639-1158	g2634058	403	423	1.10E-39	50	38	(Z99112) similar to processing protease [Bacillus subtilis]
749	Bt1Gc1054	Bt1G1009	474-1	g1805443	459	471	9.30E-45	58	69	(D50453) homologue of hypothetical protein in a rapamycin synthesis gene cluster of Streptomyces hygroscopicus [Bacillus subtilis]
749	Bt1Gc1054	Bt1G1010	1919-492	g1805444	834	678	1.10E-66	39	100	(D50453) homologue of hypothetical protein in a rapamycin synthesis gene cluster of Streptomyces hygroscopicus [Bacillus subtilis]
750	Bt1Gc1052	Bt1G1011	803-1	g1084221	1278	1282	1.10E-130	95	86	beta-lactamase (EC 3.5.2.6) precursor - Bacillus thuringiensis [Bacillus thuringiensis]
750	Bt1Gc1052	Bt1G1012	2535-609	g1731017	1198	1219	5.10E-124	52	71	HYPOTHETICAL 80.1 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
751	Bt1Gc1057	Bt1G1013	348-1	g729343	286	291	1.10E-25	53	29	PROTEIN DLT.D PRECURSOR [Bacillus subtilis]
751	Bt1Gc1057	Bt1G1014	2153-647	g2633558	110	166	2.00E-12	34	66	(Z99110) similar to hypothetical proteins [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
751	Bt1Gc1057	Bt1G1015	2633-1275	g2632017	745	830	8.50E-83	38	82	(AJ002571) DppE [Bacillus subtilis]
752	Bt1Gc1058	Bt1G1016	198-1	g3282095	226	227	6.70E-19	59	23	(AJ007446) hypothetical protein [Thermotoga neapolitana]
752	Bt1Gc1058	Bt1G1017	1034-204	g1709566	992	992	5.80E-100	71	100	3-METHYL-2-OXOBUTANOATE HYDROXYMETHYLTRANSFERASE (KETOPANTOATE)
753	Bt1Gc1059	Bt1G1018	164-2198	g1731017	1593	1725	1.20E-177	53	93	HYDROXYMETHYLTRANSFERASE [Bacillus subtilis]
754	Bt1Gc1061	Bt1G1019	78-1394	g3023410	1071	863	2.70E-86	50	100	HYPOTHETICAL 80.1 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
754	Bt1Gc1061	Bt1G1020	1481-1807	g2851549	382	394	1.30E-36	71	99	BRANCHED-CHAIN AMINO ACID TRANSPORT SYSTEM CARRIER PROTEIN (BRANCHED-CHAIN AMINO ACID UPTAKE CARRIER) [Bacillus subtilis]
755	Bt1Gc1063	Bt1G1021	1273-750	g773349	477	395	1.10E-36	56	53	CSAA PROTEIN [Bacillus subtilis]
756	Bt1Gc1062	Bt1G1022	1499-2618	g1788760	159	131	2.00E-08	49	32	(U20445) BirA protein [Bacillus subtilis]
757	Bt1Gc1064	Bt1G1023	1021-221	g1881234	447	414	1.00E-38	35	100	(AE000329) orf, hypothetical protein [Escherichia coli]
757	Bt1Gc1064	Bt1G1024	1446-1653	g1722938	129	137	2.30E-09	39	51	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
758	Bt1Gc1065	Bt1G1025	728-186	g2634575	133	191	4.40E-15	26	99	HYPOTHETICAL 14.3 KD PROTEIN CY427.14 [Mycobacterium tuberculosis]
759	Bt1Gc1066	Bt1G1026	1-362	g3913081	469	472	7.30E-45	71	81	(Z99115) similar to phage-related protein [Bacteriophage SPBc2]
759	Bt1Gc1066	Bt1G1027	630-1962	g132256	1200	1187	1.30E-120	56	77	ARGININE REPRESSOR [Bacillus stearothermophilus]
760	Bt1Gc1070	Bt1G1028	453-950	g4584100	806	811	8.70E-81	91	99	DNA REPAIR PROTEIN RECN (RECOMBINATION PROTEIN N) [Bacillus subtilis]
761	Bt1Gc1069	Bt1G1029	1699-1	g1708643	439	453	9.10E-43	33	52	(AJ010132) hypothetical protein [Bacillus cereus]
761	Bt1Gc1069	Bt1G1030	1821-1205	g2829805	330	371	3.70E-34	39	75	(U51115) YebA [Bacillus subtilis]
762	Bt1Gc1071	Bt1G1031	1-1083	g2626818	874	675	2.30E-66	46	92	HYPOTHETICAL 30.5 KD PROTEIN IN GABP-GUAA INTERGENIC REGION [Bacillus subtilis]
762	Bt1Gc1071	Bt1G1032	1088-1282	g2239294	195	200	4.90E-16	58	98	(D83967) YrkF [Bacillus subtilis]
763	Bt1Gc1074	Bt1G1033	1-704	g1001758	60	122	1.00E-07	33	37	(U51115) unknown [Bacillus subtilis]
764	Bt1Gc1072	Bt1G1034	863-1135	g2622856	228	243	1.40E-20	51	99	(D64004) hypothetical protein [Synechocystis sp.]
764	Bt1Gc1072	Bt1G1035	1466-2299	g1783243	1031	753	1.20E-74	69	100	(AE000929) unknown [Methanobacterium thermoautotrophicum]
765	Bt1Gc1073	Bt1G1036	1001-1	g135566	61	145	3.30E-09	29	21	(D83026) homologous to jojC gene product (B. subtilis; prf211327a); hypothetical [Bacillus subtilis]
										TRYPSIN-RESISTANT SURFACE T6 PROTEIN PRECURSOR [Streptococcus pyogenes]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
765	Bt1Gc1073	Bt1G1037	1001-1	g135566	61	145	3.30E-09	29	21	TRYPSIN-RESISTANT SURFACE T6 PROTEIN PRECURSOR [Streptococcus pyogenes]
765	Bt1Gc1073	Bt1G1038	2548-838	g3036999	212	253	1.20E-21	45	28	(AF019629) putative fimbria-associated protein [Actinomyces naeslundii]
766	Bt1Gc1068	Bt1G1039	1-446	g134649	501	504	3.00E-48	58	73	SUPEROXIDE DISMUTASE (MN) [Bacillus caldotenax]
766	Bt1Gc1068	Bt1G1040	2220-1	g2765081	44	163	1.90E-11	29	40	(Y10557) g5bf [Arabidopsis thaliana]
766	Bt1Gc1068	Bt1G1041	2125-742	g2498126	320	367	9.80E-34	36	100	PUTATIVE TRANSPORTER PROTEIN AMIS2 [Rhodococcus sp. R312]
767	Bt1Gc1075	Bt1G1042	633-290	g1786884	59	127	2.60E-08	27	99	(AE000170) putative RNA [Escherichia coli]
768	Bt1Gc1076	Bt1G1043	1190-1909	g2633540	664	664	3.30E-65	52	100	(Z99110) yjcH [Bacillus subtilis]
769	Bt1Gc1077	Bt1G1044	1-529	g4539176	195	199	6.20E-16	31	61	(AL049485) hypothetical protein [Streptomyces coelicolor]
770	Bt1Gc1078	Bt1G1045	1-1614	g121408	1837	1719	5.30E-177	66	97	AEROBIC GLYCEROL-3-PHOSPHATE DEHYDROGENASE [Bacillus subtilis]
771	Bt1Gc1079	Bt1G1046	1-1100	g2443219	739	605	5.90E-59	41	72	(D86417) YfkQ [Bacillus subtilis]
771	Bt1Gc1079	Bt1G1047	1078-1984	g3290176	310	318	1.50E-28	27	82	(AF067645) spore germination protein GerB [Bacillus cereus]
772	Bt1Gc1082	Bt1G1048	910-1	g1730894	1326	1355	2.00E-138	84	72	HYPOTHETICAL 46.7 KD OXIDOREDUCTASE IN RECQ-CMK INTERGENIC REGION [Bacillus subtilis]
772	Bt1Gc1082	Bt1G1049	1843-1256	g1730890	601	645	3.40E-63	61	99	HYPOTHETICAL 22.2 KD PROTEIN IN RECQ-CMK INTERGENIC REGION [Bacillus subtilis]
773	Bt1Gc1080	Bt1G1050	1554-70	g2619016	1652	1656	2.50E-170	61	100	(AF027868) aldehyde dehydrogenase [Bacillus subtilis]
773	Bt1Gc1080	Bt1G1051	2060-1575	g2983599	126	171	1.10E-12	25	54	(AE000725) dihydrodipicolinate synthase [Aquifex aeolicus]
774	Bt1Gc1083	Bt1G1052	1-557	g1717743	422	430	2.10E-40	49	78	TREHALOSE OPERON TRANSCRIPTIONAL REPRESSOR [Bacillus subtilis]
774	Bt1Gc1083	Bt1G1053	704-2122	g2443218	1255	1163	4.40E-118	52	100	(D86417) TreP [Bacillus subtilis]
775	Bt1Gc1084	Bt1G1054	180-1484	g730002	974	871	3.80E-87	44	66	METHYL-ACCEPTING CHEMOTAXIS PROTEIN MCPA (H1) [Bacillus subtilis]
776	Bt1Gc1086	Bt1G1055	651-163	g1405451	477	374	1.80E-34	58	99	(Z73234) YneJ [Bacillus subtilis]
776	Bt1Gc1086	Bt1G1056	1597-1190	g116041	442	448	2.60E-42	61	99	CYTIDINE DEAMINASE (CYTIDINE AMINOHYDROLASE) (CDA) [Bacillus subtilis]
776	Bt1Gc1086	Bt1G1057	2420-1634	g730287	1083	1085	8.10E-110	84	60	PYRIMIDINE-NUCLEOSIDE PHOSPHORYLASE [Bacillus subtilis]
777	Bt1Gc1087	Bt1G1058	1-625	g1731002	518	537	9.50E-52	50	97	HYPOTHETICAL 23.7 KD PROTEIN IN CCCA-SODA INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
777	Bt1Gc1087	Bt1G1059	619-1738	g2634950	1140	1103	1.00E-111	59	100	(Z99116) similar to hypothetical proteins [Bacillus subtilis]
778	Bt1Gc1088	Bt1G1060	418-1	g2499952	636	638	1.90E-62	86	67	URACIL PHOSPHORIBOSYLTRANSFERASE (UMP PYROPHOSPHORYLASE) (UPRTASE) []
778	Bt1Gc1088	Bt1G1061	1932-688	g729608	1744	1752	1.70E-180	80	100	SERINE HYDROXYMETHYLTRANSFERASE (SERINE METHYLASE) (SHMT) [Bacillus subtilis]
779	Bt1Gc1085	Bt1G1062	740-1648	g1724002	762	732	2.10E-72	50	100	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN 2 IN GLPD-CSPB INTERGENIC REGION [Bacillus subtilis]
779	Bt1Gc1085	Bt1G1063	1635-2579	g1724003	372	414	1.00E-38	30	100	HYPOTHETICAL 34.9 KD PROTEIN IN GLPD-CSPB INTERGENIC REGION [Bacillus subtilis]
780	Bt1Gc1089	Bt1G1064	201-794	g3308998	473	516	1.60E-49	51	99	(AB015998) CBP21 precursor [Serratia marcescens]
780	Bt1Gc1089	Bt1G1065	1-1546	g1708084	293	431	2.50E-39	34	24	EXOGLUCANASE B PRECURSOR (EXOCHELLOBIOHYDROLASE B) (1,4-BETA-CELLOBIOHYDROLASE B) (CBP120) [Cellulomonas fimi]
781	Bt1Gc1090	Bt1G1066	1-1484	g1881239	140	228	4.20E-18	33	23	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
781	Bt1Gc1090	Bt1G1067	1305-550	g1176064	427	473	5.70E-45	37	100	HYPOTHETICAL PROTEIN H11664 []
781	Bt1Gc1090	Bt1G1068	2300-2003	g1789981	276	299	1.60E-26	54	35	(AE000433) IS150 putative transposase [Escherichia coli]
782	Bt1Gc1091	Bt1G1069	634-188	g1730120	643	643	5.50E-63	81	99	FERRIC UPTAKE REGULATION PROTEIN HOMOLOG 2 [Bacillus subtilis]
782	Bt1Gc1091	Bt1G1070	1390-752	g586019	518	381	3.20E-35	48	100	STAGE II SPORULATION PROTEIN M [Bacillus subtilis]
783	Bt1Gc1094	Bt1G1071	757-198	g2280497	100	124	2.50E-07	26	99	(AB005554) probable glucose 1-dehydrogenase [Bacillus subtilis]
784	Bt1Gc1096	Bt1G1072	294-1173	g584979	207	152	5.90E-11	37	99	SPORE COAT PROTEIN X [Bacillus subtilis]
785	Bt1Gc1093	Bt1G1073	385-1	g1001505	208	234	1.20E-19	36	37	(D64000) ABC transporter [Synechocystis sp.]
785	Bt1Gc1093	Bt1G1074	1785-1	g2293324	259	321	7.30E-29	34	68	(AF008220) LysR family transcription regulator [Bacillus subtilis]
785	Bt1Gc1093	Bt1G1075	1896-2252	g2635785	176	222	2.30E-18	40	90	(Z99120) similar to acylolate catabolism [Bacillus subtilis]
786	Bt1Gc1092	Bt1G1076	311-1	g586169	200	185	1.90E-14	41	37	UREASE ACCESSORY PROTEIN URED [Bacillus sp.]
786	Bt1Gc1092	Bt1G1077	925-314	g586172	782	709	5.60E-70	74	100	UREASE ACCESSORY PROTEIN UREG [Bacillus sp.]
786	Bt1Gc1092	Bt1G1078	1626-946	g586171	471	473	5.70E-45	42	100	UREASE ACCESSORY PROTEIN UREF [Bacillus sp.]
786	Bt1Gc1092	Bt1G1079	2065-1622	g586170	310	321	7.30E-29	40	99	UREASE ACCESSORY PROTEIN UREE [Bacillus sp.]
787	Bt1Gc1097	Bt1G1080	1151-1	g1168442	157	300	1.20E-26	29	64	BETA-LACTAMASE PRECURSOR (CEPHALOSPORINASE) [Yersinia enterocolitica]
787	Bt1Gc1097	Bt1G1081	1605-1431	g2226127	261	280	1.60E-24	88	18	(Y14078) Hypothetical protein [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
788	Bt1Gc1098	Bt1G1082	2536-745	g2507582	192	228	1.40E-17	31	17	HYPOTHETICAL 138.1 KD PROTEIN IN MOLR-BGLX INTERGENIC REGION [Escherichia coli]
789	Bt1Gc1099	Bt1G1083	699-1	g2634067	211	295	4.20E-26	32	51	(Z99112) penicillin-binding protein [Bacillus subtilis]
790	Bt1Gc1100	Bt1G1084	782-1	g72992	991	893	1.80E-89	73	75	recE protein - Bacillus subtilis []
791	Bt1Gc1102	Bt1G1085	1-1250	g2688588	493	553	1.00E-55	33	79	(AE001167) lysyl-tRNA synthetase [Borrelia burgdorferi]
791	Bt1Gc1102	Bt1G1086	1740-1311	g1175629	434	455	4.60E-43	57	99	CELL WALL HYDROLASE CWLJ [Bacillus subtilis]
792	Bt1Gc1105	Bt1G1087	608-93	g2632232	469	478	1.70E-45	52	99	(AJ222587) YkuK protein [Bacillus subtilis]
792	Bt1Gc1105	Bt1G1088	1310-1077	g2632231	247	256	5.70E-22	62	99	(AJ222587) YkuJ protein [Bacillus subtilis]
793	Bt1Gc1106	Bt1G1089	942-1	g4155714	382	418	3.90E-39	35	72	(AE001539) HISTIDYL-TRNA SYNTHETASE [Helicobacter pylori J99]
793	Bt1Gc1106	Bt1G1090	1274-1750	g2633560	435	438	2.90E-41	51	99	(Z99110) similar to transcription regulation [Bacillus subtilis]
794	Bt1Gc1095	Bt1G1091	1-451	g2633543	116	158	1.40E-11	26	83	(Z99110) similar to ribosomal-protein-alanine N-acetyltransferase [Bacillus subtilis]
794	Bt1Gc1095	Bt1G1092	2172-860	g3258354	943	827	1.80E-82	49	100	(AP000007) 424aa long hypothetical protein [Pyrococcus horikoshii]
795	Bt1Gc1101	Bt1G1093	530-1	g1730265	491	486	2.40E-46	58	99	HYPOTHETICAL 20.1 KD PROTEIN IN HMP 5'REGION (ORF1) [Bacillus subtilis]
796	Bt1Gc1110	Bt1G1094	289-1	g2619009	95	139	1.50E-09	28	43	(AF027868) RAS-related protein [Bacillus subtilis]
796	Bt1Gc1110	Bt1G1095	881-699	g1731301	89	132	7.80E-09	45	51	HYPOTHETICAL 13.0 KD PROTEIN IN IDH-DEOR INTERGENIC REGION PRECURSOR [Bacillus subtilis]
796	Bt1Gc1110	Bt1G1096	2084-2868	g67770	1059	1057	7.50E-107	86	81	beta-lactamase (EC 3.5.2.6) III precursor - Bacillus cereus []
797	Bt1Gc1111	Bt1G1097	985-1275	g4894353	321	321	7.30E-29	64	99	(AF065404) pXO1-138 [Bacillus anthracis]
798	Bt1Gc1109	Bt1G1098	1-1452	g3914289	2092	1955	5.20E-202	83	74	TOPOISOMERASE IV SUBUNIT B [Bacillus subtilis]
798	Bt1Gc1109	Bt1G1099	1454-1695	g2558947	324	294	6.50E-25	79	10	(AF024713) ParC [Bacillus subtilis]
799	Bt1Gc1108	Bt1G1100	Jan-66	g2497392	101	101	5.10E-05	91	8	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
799	Bt1Gc1108	Bt1G1101	944-309	g79959	243	326	2.20E-29	37	100	hypothetical 23.9K protein - Enterococcus faecalis plasmid pAM-beta-1 [Enterococcus faecalis]
799	Bt1Gc1108	Bt1G1102	1192-2078	g1694898	358	411	2.80E-37	30	29	(Y09450) transposase [Pseudomonas putida]
800	Bt1Gc1112	Bt1G1103	563-230	g2635198	349	366	1.20E-33	68	99	(Z99117) similar to hypothetical proteins [Bacillus subtilis]
800	Bt1Gc1112	Bt1G1104	857-1738	g2635199	1063	1064	1.40E-107	71	70	(Z99117) similar to hypothetical proteins [Bacillus subtilis]
801	Bt1Gc1113	Bt1G1105	1-434	g2415743	431	276	4.30E-24	59	56	(AB000617) YceF [Bacillus subtilis]
801	Bt1Gc1113	Bt1G1106	550-1819	g2415744	847	861	4.40E-86	41	78	(AB000617) YceG [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
802	Bt1Gc1115	Bt1G1107	644-1195	g2266414	681	681	5.20E-67	70	99	(Y10549) N-formylmethionylaminoacyl-tRNA deformylase [Bacillus stearothermophilus]
802	Bt1Gc1115	Bt1G1108	1790-1259	g2633826	427	452	9.60E-43	48	68	(Z99111) similar to hypothetical proteins [Bacillus subtilis]
803	Bt1Gc1119	Bt1G1109	1053-1	g3688818	744	600	3.40E-71	45	95	(AF084104) hypothetical protein [Bacillus firmus]
804	Bt1Gc1116	Bt1G1110	1637-1	g4894306	480	612	1.10E-59	43	42	(AF065404) pXOI-90 [Bacillus anthracis]
805	Bt1Gc1118	Bt1G1111	1-1457	g2833392	1119	1184	2.60E-120	47	87	HYPOTHETICAL 61.5 KD PROTEIN IN ADEC-PDHA INTERGENIC REGION [Bacillus subtilis]
806	Bt1Gc1117	Bt1G1112	25-738	g141397	846	758	3.60E-75	70	100	HYPOTHETICAL OXIDOREDUCTASE IN RTP-PELB INTERGENIC REGION (ORF238) [Bacillus subtilis]
806	Bt1Gc1117	Bt1G1113	1032-1790	g112961	541	501	6.20E-48	42	89	AMINOGLYCOSIDE 6-ADENYLYLTRANSFERASE (AAD(6)) [Bacillus subtilis]
807	Bt1Gc1125	Bt1G1114	1-1310	g1881343	687	549	5.10E-53	50	87	(AB001488) FUNCTION UNKNOWN, SIMILAR PRODUCT IN E. COLI. [Bacillus subtilis]
807	Bt1Gc1125	Bt1G1115	932-1416	g121839	540	555	1.20E-53	65	37	ENDOGLUCANASE PRECURSOR (ENDO-1,4-BETA-GLUCANASE) (CELLULASE) (ENDO-K) [Bacillus sp.]
808	Bt1Gc1122	Bt1G1116	510-1	g2635805	521	535	1.50E-51	61	81	(Z99120) similar to two-component response regulator [Bacillus subtilis]
808	Bt1Gc1122	Bt1G1117	1562-487	g2635806	704	634	5.00E-62	45	100	(Z99120) similar to two-component sensor histidine kinase [Bacillus subtilis]
808	Bt1Gc1122	Bt1G1118	2289-1565	g2635807	286	218	6.00E-18	31	100	(Z99120) yvqF [Bacillus subtilis]
809	Bt1Gc1120	Bt1G1119	1616-215	g1731094	1894	1603	1.00E-164	78	100	HYPOTHETICAL NA+/H+ ANTIPORTER IN ANSB-SPOIIM INTERGENIC REGION [Bacillus subtilis]
810	Bt1Gc1124	Bt1G1120	1-1041	g2337813	1000	839	9.40E-84	57	62	(Y13937) YloV protein [Bacillus subtilis]
810	Bt1Gc1124	Bt1G1121	1335-1941	g3914611	623	637	2.40E-62	58	30	ATP-DEPENDENT DNA HELICASE RECG [Bacillus subtilis]
811	Bt1Gc1121	Bt1G1122	1437-350	g3328828	675	527	1.10E-50	38	89	(AE001313) Glutamate Symport [Chlamydia trachomatis]
812	Bt1Gc1126	Bt1G1123	1385-1	g586703	283	239	8.40E-20	46	33	HYPOTHETICAL 43.0 KD PROTEIN IN PROK-TAG INTERGENIC REGION (F402) [Escherichia coli]
812	Bt1Gc1126	Bt1G1124	1599-870	g2499154	503	514	2.60E-49	42	71	HYPOTHETICAL 37.7 KD PROTEIN IN GRPA 5'REGION (ORFA) [Bacillus firmus]
813	Bt1Gc1123	Bt1G1125	970-1	g3122129	1287	1223	1.90E-124	73	95	GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE 2 (GAPDH) [Bacillus subtilis]
813	Bt1Gc1123	Bt1G1126	1644-1089	g3183497	480	507	1.40E-48	56	94	HYPOTHETICAL 22.0 KD PROTEIN IN GAPB-MUTM INTERGENIC REGION [Bacillus subtilis]
814	Bt1Gc1130	Bt1G1127	503-225	g2128793	115	159	1.10E-11	36	99	hypothetical protein MJ1325 - Methanococcus jannaschii

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
814	Bt1Gc1130	Bt1G1128	1026-1942	g1076106	369	379	6.40E-35	32	54	[Methanococcus jannaschii]
815	Bt1Gc1132	Bt1G1129	1159-938	g2649348	71	143	5.30E-10	38	53	schC protein - Streptomyces halstedii []
815	Bt1Gc1132	Bt1G1130	1435-1	g2634950	84	130	7.50E-08	33	23	(AE001019) mutator protein MutT (mutT) [Archaeoglobus fulgidus]
816	Bt1Gc1133	Bt1G1131	1-1026	g1708267	842	891	2.90E-89	53	84	FLAVOHEMOPROTEIN (HAEMOGLOBIN-LIKE PROTEIN) (FLAVOHEMOGLOBIN) [Bacillus subtilis]
817	Bt1Gc1128	Bt1G1132	1-855	g2465008	344	471	9.30E-45	42	77	(AJ001445) ripening-induced protein [Fragaria vesca]
818	Bt1Gc1131	Bt1G1133	147-1600	g2493592	647	572	1.90E-55	40	100	HYPOTHETICAL 43.7 KD PROTEIN IN NPPE-PYCA INTERGENIC REGION [Bacillus subtilis]
818	Bt1Gc1131	Bt1G1134	859-1963	g3913213	150	245	2.50E-20	28	43	3-CHLOROBENZATE-3,4-DIOXYGENASE OXYGENASE SUBUNIT [Alcaligenes sp.]
819	Bt1Gc1129	Bt1G1135	324-2325	g225714	1796	1388	7.40E-167	54	86	urease [Canavalia ensiformis]
820	Bt1Gc1135	Bt1G1136	1859-1	g1213021	578	477	2.20E-45	44	63	(X84710) ORF492, surface antigen gene [Methanosarcina mazei]
821	Bt1Gc1136	Bt1G1137	306-1	g3257514	146	142	4.30E-09	34	25	(AP000005) 403aa long hypothetical macrolide-efflux determinant [Pyrococcus horikoshii]
821	Bt1Gc1136	Bt1G1138	1746-1922	g2633135	132	135	3.80E-09	47	98	(Z99108) yfU [Bacillus subtilis]
822	Bt1Gc1140	Bt1G1139	1018-1	g3724051	813	773	9.30E-77	46	91	(AJ000758) hypothetical protein [Bacillus megaterium]
823	Bt1Gc1143	Bt1G1140	1-273	g1172978	364	364	2.00E-33	79	75	50S RIBOSOMAL PROTEIN L18 [Bacillus subtilis]
823	Bt1Gc1143	Bt1G1141	298-795	g1173265	748	748	4.10E-74	90	99	30S RIBOSOMAL PROTEIN S5 (BS5) [Bacillus subtilis]
823	Bt1Gc1143	Bt1G1142	812-997	g4512422	221	226	8.60E-19	73	98	(AB017508) rpmD homologue (identity of 82% to B. subtilis) [Bacillus halodurans]
823	Bt1Gc1143	Bt1G1143	1028-1139	g4512423	158	96	5.10E-05	95	25	(AB017508) rplO homologue (identity of 84% to B. subtilis) [Bacillus halodurans]
824	Bt1Gc1138	Bt1G1144	775-1	g3688825	638	677	1.40E-66	51	88	(AF084104) NatA [Bacillus firmus]
824	Bt1Gc1138	Bt1G1145	1103-1609	g1944618	285	327	1.70E-29	38	61	(Y12602) acid phosphatase [Streptococcus equisimilis]
825	Bt1Gc1142	Bt1G1146	1017-1	g2851670	952	850	6.40E-85	56	78	HYPOTHETICAL 49.0 KD PROTEIN IN BLTD-TRKA INTERGENIC REGION [Bacillus subtilis]
826	Bt1Gc1134	Bt1G1147	1-369	g2326729	155	176	1.70E-13	32	54	(Z98268) hypothetical protein Rv1686c [Mycobacterium tuberculosis]
826	Bt1Gc1134	Bt1G1148	442-1047	g1770039	126	169	9.40E-13	24	99	(Z75208) hypothetical protein [Bacillus subtilis]
826	Bt1Gc1134	Bt1G1149	1709-1872	g2621999	128	145	1.20E-09	47	17	(AE000866) photoreactivation-associated protein [Methanobacterium thermoautotrophicum]
827	Bt1Gc1141	Bt1G1150	387-749	g2633753	294	358	8.80E-33	56	58	(Z99111) similar to spore cortex-lytic enzyme [Bacillus

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
828	Bt1Gc1145	Bt1G1151	479-39	g2632544	317	285	4.80E-25	41	99	(Z99105) ycbO [Bacillus subtilis]
828	Bt1Gc1145	Bt1G1152	2115-1675	g2632544	214	160	8.40E-12	33	99	(Z99105) ycbO [Bacillus subtilis]
829	Bt1Gc1144	Bt1G1153	2448-1366	g132565	946	982	6.60E-99	52	100	RIBOFLAVIN-SPECIFIC DEAMINASE [Bacillus subtilis]
830	Bt1Gc1150	Bt1G1154	230-988	g2117240	334	377	8.50E-35	34	100	(Z95586) hypothetical protein Rv1597 [Mycobacterium tuberculosis]
830	Bt1Gc1150	Bt1G1155	1696-1216	g730088	118	207	8.80E-17	34	99	MUTATOR MUTT PROTEIN (7,8-DIHYDRO-8-OXOGUANINE-TRIPHOSPHATASE) (8-OXO-DGTPASE) (DGTP PYROPHOSPHOHYDROLASE) [Streptococcus pneumoniae]
830	Bt1Gc1150	Bt1G1156	2299-1851	g585393	311	332	5.00E-30	43	82	SIGNAL PEPTIDASE I P (SPASE I) (LEADER PEPTIDASE I) [Bacillus subtilis]
831	Bt1Gc1154	Bt1G1157	383-1	g4154045	140	189	9.20E-15	34	43	(AL035159) putative CDP-diacylglycerol--serine O-phosphatidyltransferase [Mycobacterium leprae]
831	Bt1Gc1154	Bt1G1158	1805-1927	g2224768	107	133	6.10E-09	60	57	(Z97025) ylaI [Bacillus subtilis]
832	Bt1Gc1153	Bt1G1159	944-1	g417314	1561	1558	6.10E-160	99	94	MBL PROTEIN [Bacillus cereus]
832	Bt1Gc1153	Bt1G1160	1476-1108	g4584136	629	629	1.70E-61	99	99	(AJ010138) stage III sporulation protein D [Bacillus cereus]
833	Bt1Gc1146	Bt1G1161	985-170	g4753870	491	514	2.60E-49	38	100	(AL049754) putative oxidoreductase [Streptomyces coelicolor]
833	Bt1Gc1146	Bt1G1162	1668-1036	g1653142	248	296	3.30E-26	33	100	(D90911) hypothetical protein [Synechocystis sp.]
833	Bt1Gc1146	Bt1G1163	2652-1704	g3913969	504	564	1.30E-54	36	71	KYNURENINASE (L-KYNURENINE HYDROLASE) [Rattus norvegicus]
834	Bt1Gc1152	Bt1G1164	990-1	g1881291	294	381	3.20E-35	29	95	(AB001488) PROBABLE INTEGRASE. [Bacillus subtilis]
834	Bt1Gc1152	Bt1G1165	993-1	g166159	321	335	2.40E-30	30	95	(M34832) integrase (int) [Bacteriophage phi-11]
834	Bt1Gc1152	Bt1G1166	1451-1011	g4126639	271	292	8.70E-26	42	99	(AB016282) ORF2 [bacteriophage phi-105]
835	Bt1Gc1147	Bt1G1167	480-31	g1881290	510	431	1.60E-40	59	99	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
835	Bt1Gc1147	Bt1G1168	1566-860	g2632778	817	706	1.20E-69	68	30	(Z99106) similar to hypothetical proteins [Bacillus subtilis]
836	Bt1Gc1148	Bt1G1169	269-1	g3930535	72	152	2.10E-09	32	7	(AF065313) DNA polymerase III; family C DNA polymerase [Thermotoga maritima]
836	Bt1Gc1148	Bt1G1170	699-502	g2493764	306	306	2.90E-27	85	98	COLD SHOCK-LIKE PROTEIN CSPD [Bacillus cereus]
836	Bt1Gc1148	Bt1G1171	1106-1816	g2649993	202	208	6.90E-17	35	100	(AE001061) conserved hypothetical protein [Archaeoglobus fulgidus]
836	Bt1Gc1148	Bt1G1172	1991-2222	g225559	360	376	1.10E-34	96	16	ORF IS231C [Bacillus thuringiensis]
837	Bt1Gc1151	Bt1G1173	1-941	g2632613	578	623	7.30E-61	41	92	(Z99105) similar to thioredoxin reductase [Bacillus subtilis]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
838	Bt1Gc1156	Bt1G1174	420-1	g3183517	449	464	5.20E-44	59	97	GENERAL STRESS PROTEIN 20U (GSP20U) (DPS PROTEIN HOMOLOG) [Bacillus subtilis]
838	Bt1Gc1156	Bt1G1175	670-1715	g2633742	607	453	7.50E-43	41	100	(Z99111) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
839	Bt1Gc1157	Bt1G1176	1-298	g1709688	330	338	1.20E-30	65	56	PUTATIVE PEPTIDE METHIONINE SULFOXIDE REDUCTASE (PEPTIDE MET(O) REDUCTASE) [Bacillus subtilis]
840	Bt1Gc1160	Bt1G1177	680-1	g2493592	302	256	9.40E-22	34	58	HYPOTHETICAL 43.7 KD PROTEIN IN NPPE-PYCA INTERGENIC REGION [Bacillus subtilis]
840	Bt1Gc1160	Bt1G1178	1090-812	g2224773	352	352	3.80E-32	73	99	(Z97025) ylaN [Bacillus subtilis]
840	Bt1Gc1160	Bt1G1179	1249-1740	g2224771	265	296	3.30E-26	37	99	(Z97025) ylaL [Bacillus subtilis]
840	Bt1Gc1160	Bt1G1180	1995-1778	g2224770	231	254	2.70E-21	71	16	(Z97025) product similar to E. coli PhoH protein [Bacillus subtilis]
841	Bt1Gc1163	Bt1G1181	1-1488	g2226166	1160	1066	4.00E-110	46	85	(Y14080) hypothetical protein [Bacillus subtilis]
842	Bt1Gc1161	Bt1G1182	160-1	g4512355	149	185	1.90E-14	69	41	(AB011836) similar to B.subtilis ywgB gene(27%-identity) [Bacillus halodurans]
842	Bt1Gc1161	Bt1G1183	2389-1016	g2245640	1402	1350	6.70E-138	57	100	(AF006075) dihydroipoamide dehydrogenase [Bacillus subtilis]
842	Bt1Gc1161	Bt1G1184	2983-2411	g2633132	544	548	6.50E-53	55	48	(Z99108) acetoin dehydrogenase E2 component (dihydroipoamide acetyltransferase) [Bacillus subtilis]
843	Bt1Gc1159	Bt1G1185	1132-221	g2633815	763	707	9.20E-70	49	100	(Z99111) similar to thiamin biosynthesis [Bacillus subtilis]
844	Bt1Gc1162	Bt1G1186	1160-189	g2612898	812	711	3.50E-70	48	100	(AF015825) YisP-like protein [Bacillus subtilis]
845	Bt1Gc1164	Bt1G1187	1-251	g2635785	156	166	2.00E-12	41	63	(Z99120) similar to acylate catabolism [Bacillus subtilis]
846	Bt1Gc1165	Bt1G1188	1-856	g3328409	314	281	2.50E-23	30	28	(AE001277) Isoleucyl-tRNA Synthetase [Chlamydia trachomatis]
846	Bt1Gc1165	Bt1G1189	1-853	g3257481	229	237	1.30E-18	29	26	(AP000004) 1066aa long hypothetical isoleucyl-tRNA synthetase [Pyrococcus horikoshii]
847	Bt1Gc1168	Bt1G1190	586-1	g4514331	135	207	8.80E-17	27	73	(AB013369) ComEC [Bacillus halodurans]
847	Bt1Gc1168	Bt1G1191	1324-744	g466186	451	492	5.60E-47	50	99	HYPOTHETICAL 22.0 KD PROTEIN IN RIBT-DACB INTERGENIC REGION (ORFX8) [Bacillus subtilis]
847	Bt1Gc1168	Bt1G1192	2223-1473	g466185	738	698	8.20E-69	58	100	HYPOTHETICAL 29.6 KD PROTEIN IN RIBT-DACB INTERGENIC REGION (ORFX7) [Bacillus subtilis]
848	Bt1Gc1171	Bt1G1193	1192-1	g1711644	244	341	8.40E-31	39	37	PROBABLE ASPARAGINYL-TRNA SYNTHETASE (ASPARAGINE--TRNA LIGASE) (ASNRS) [Synecocystis sp.]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
848	Bt1Gc1171	Bt1G1194	1570-881	g3123292	668	605	5.90E-59	55	28	PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PHERS) [Bacillus subtilis]
849	Bt1Gc1172	Bt1G1195	Jan-88	g2735510	97	106	4.50E-06	67	53	(U96108) (3R)-hydroxymyristoyl acyl carrier protein dehydrase homolog [Staphylococcus carnosus]
849	Bt1Gc1172	Bt1G1196	439-1116	g1495280	568	599	2.50E-58	51	100	(Z71928) hypothetical protein [Bacillus subtilis]
850	Bt1Gc1173	Bt1G1197	1856-1	g730965	2463	2273	3.50E-237	76	89	DNA TOPOISOMERASE I (OMEGA-PROTEIN) (RELAXING ENZYME) (UNTWISTING ENZYME) (SWIVELASE) [Bacillus subtilis]
850	Bt1Gc1173	Bt1G1198	2540-1293	g3915864	407	428	3.40E-40	48	63	SMF PROTEIN [Bacillus subtilis]
851	Bt1Gc1169	Bt1G1199	198-1	g448837	203	211	3.30E-17	59	32	superoxide dismutase [Bacillus stearothermophilus]
851	Bt1Gc1169	Bt1G1200	500-1501	g2443235	1052	1017	1.30E-102	58	100	(D86417) YfM [Bacillus subtilis]
852	Bt1Gc1175	Bt1G1201	351-34	g1881265	376	380	4.10E-35	58	99	(AB001488) PROBABLE THIOREDOXIN. [Bacillus subtilis]
852	Bt1Gc1175	Bt1G1202	2147-609	g141186	72	123	7.00E-08	42	37	HYPOTHETICAL P20 PROTEIN [Bacillus licheniformis]
853	Bt1Gc1178	Bt1G1203	1-418	g1405459	121	131	1.00E-08	27	72	(Z73234) YneS [Bacillus subtilis]
853	Bt1Gc1178	Bt1G1204	984-2096	g2649155	155	250	2.40E-21	32	70	(AE001006) aspartate aminotransferase (aspC) [Archaeoglobus fulgidus]
854	Bt1Gc1176	Bt1G1205	857-1	g2633600	484	487	1.90E-46	53	59	(Z99110) N-acetylmuramoyl-L-alanine amidase [Bacillus subtilis]
854	Bt1Gc1176	Bt1G1206	1276-884	g141088	185	194	2.10E-15	34	99	HYPOTHETICAL 14.9 KD PROTEIN IN NAGH 3'REGION (ORFD) [Clostridium perfringens]
855	Bt1Gc1177	Bt1G1207	1634-417	g2226187	812	769	2.50E-76	40	100	(Y14081) hypothetical protein [Bacillus subtilis]
855	Bt1Gc1177	Bt1G1208	1724-3056	g2226188	678	759	2.80E-75	37	76	(Y14081) hypothetical protein [Bacillus subtilis]
856	Bt1Gc1182	Bt1G1209	815-1545	g3845215	48	74	0.72	26	2	(AE001402) hypothetical protein [Plasmodium falciparum]
857	Bt1Gc1183	Bt1G1210	40-1492	g2635246	1379	1230	3.50E-125	56	100	(Z99118) similar to sodium/proton-dependent alanine carrier protein [Bacillus subtilis]
857	Bt1Gc1183	Bt1G1211	1617-1993	g2635768	358	306	2.90E-27	55	99	(Z99120) yurZ [Bacillus subtilis]
857	Bt1Gc1183	Bt1G1212	2404-2210	g2337807	70	118	9.90E-07	34	22	(Y13937) YloQ protein [Bacillus subtilis]
858	Bt1Gc1180	Bt1G1213	1-641	g2635801	855	872	3.00E-87	78	46	(Z99120) fumarate hydratase [Bacillus subtilis]
859	Bt1Gc1184	Bt1G1214	2223-695	g3258180	537	702	3.10E-69	35	100	(AP000007) 482aa long hypothetical D-nopaline dehydrogenase [Pyrococcus horikoshii]
860	Bt1Gc1185	Bt1G1215	373-1	g4584100	211	228	5.30E-19	39	74	(AJ010132) hypothetical protein [Bacillus cereus]
860	Bt1Gc1185	Bt1G1216	1568-674	g121839	1279	1268	3.30E-129	80	64	ENDOGLUCANASE PRECURSOR (ENDO-1,4-BETA-GLUCANASE) (CELLULASE) (ENDO-K) [Bacillus sp.]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
861	Bt1Gc1186	Bt1G1217	1-360	g2500507	482	482	6.40E-46	73	70	MENAQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT (RIESKE IRON-SULFUR PROTEIN) [Bacillus stearothermophilus]
861	Bt1Gc1186	Bt1G1218	367-1037	g1168648	1120	1131	1.10E-114	94	100	MENAQUINOL-CYTOCHROME C REDUCTASE CYTOCHROME B SUBUNIT [Bacillus subtilis]
861	Bt1Gc1186	Bt1G1219	1084-1848	g1168649	1033	961	1.10E-96	73	100	MENAQUINOL-CYTOCHROME C REDUCTASE CYTOCHROME B/C SUBUNIT [Bacillus subtilis]
862	Bt1Gc1188	Bt1G1220	70-1335	g2827439	1774	1774	7.80E-183	78	100	(AF043609) aluminum resistance protein [Arthrobacter viscosus]
862	Bt1Gc1188	Bt1G1221	1526-1909	g140715	581	548	6.50E-53	95	99	REGULATORY PROTEIN GLNR [Bacillus cereus]
863	Bt1Gc1192	Bt1G1222	1413-185	g1800301	271	275	5.50E-24	23	100	(U83667) macrolide-efflux determinant [Streptococcus pneumoniae]
863	Bt1Gc1192	Bt1G1223	2109-1564	g2982980	127	161	6.60E-12	22	99	(AE000682) transcriptional regulator (TetR/AcrR family) [Aquifex aeolicus]
864	Bt1Gc1189	Bt1G1224	1-151	g1945675	124	150	7.10E-10	53	11	(Z94043) hypothetical protein [Bacillus subtilis]
865	Bt1Gc1190	Bt1G1225	688-2037	g1303913	1932	1833	4.40E-189	80	100	(D84432) YqhX [Bacillus subtilis]
866	Bt1Gc1193	Bt1G1226	1689-1	g234795	320	382	2.50E-35	31	69	(AF078135) unknown [Leptospira borgpetersenii]
866	Bt1Gc1193	Bt1G1227	1608-721	g2116764	338	414	1.00E-38	30	100	(D86418) YnfF [Bacillus subtilis]
867	Bt1Gc1191	Bt1G1228	25-807	g1170788	496	491	7.10E-47	43	100	OUTER MEMBRANE LIPOPROTEIN 3 PRECURSOR (PLP3) [Pasteurella haemolytica]
867	Bt1Gc1191	Bt1G1229	823-1210	g2506097	355	370	4.70E-34	57	38	ATP-BINDING PROTEIN ABC [Escherichia coli]
868	Bt1Gc1195	Bt1G1230	1-472	g2633171	324	363	2.60E-33	47	52	(Z99108) similar to hypothetical proteins [Bacillus subtilis]
868	Bt1Gc1195	Bt1G1231	780-1594	g2495571	95	222	2.30E-18	25	100	HYPOTHETICAL PROTEIN HI0105 [Haemophilus influenzae Rd]
869	Bt1Gc1196	Bt1G1232	452-2266	g586060	2374	2311	9.80E-240	74	92	METHIONYL-TRNA SYNTHETASE (METHIONINE--TRNA LIGASE) (METRS) [Bacillus subtilis]
870	Bt1Gc1198	Bt1G1233	1-682	g1002992	208	276	4.30E-24	43	30	(U34772) ORF375 [Dichelobacter nodosus]
870	Bt1Gc1198	Bt1G1234	1001-1291	g2160199	111	122	9.00E-08	34	99	(D13377) XpaF1 protein [Bacillus licheniformis]
871	Bt1Gc1200	Bt1G1235	584-1735	g1730958	1427	1400	3.40E-143	69	100	HYPOTHETICAL 43.5 KD PROTEIN IN COTD-KDUD INTERGENIC REGION PRECURSOR [Bacillus subtilis]
872	Bt1Gc1199	Bt1G1236	1-830	g4589056	1132	1047	8.60E-106	78	54	(AF120090) 2,3-bisphosphoglycerate-independent phosphoglycerate mutase [Bacillus megaterium]
872	Bt1Gc1199	Bt1G1237	864-1812	g2635903	1282	1251	2.10E-127	78	73	(Z99121) enolase [Bacillus subtilis]
873	Bt1Gc1197	Bt1G1238	1-1516	g2815318	718	738	4.80E-73	38	58	(AL021529) putative export protein [Streptomyces coelicolor]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
873	Bt1Gc1197	Bt1G1239	1993-2142	g1345692	142	156	2.20E-11	53	23	CHLORAMPHENICOL ACETYLTRANSFERASE (CAT) [Clostridium butyricum]
874	Bt1Gc1174	Bt1G1240	207-1	g1730937	167	176	1.70E-13	44	35	HYPOTHETICAL 21.6 KD PROTEIN IN ILVA 3'REGION [Bacillus subtilis]
874	Bt1Gc1174	Bt1G1241	724-1475	g3559998	550	623	7.30E-61	45	74	(AL031515) hypothetical protein SC5C7.08 [Streptomyces coelicolor]
874	Bt1Gc1174	Bt1G1242	2186-1862	g1620923	122	142	1.80E-09	30	39	(Z79580) putative ORF [Bacillus subtilis]
875	Bt1Gc1202	Bt1G1243	96-1016	g1894757	923	902	2.00E-90	58	100	(Z92953) periplasmic substrate-binding protein [Bacillus subtilis]
875	Bt1Gc1202	Bt1G1244	1057-1257	g141471	241	250	2.40E-21	74	36	TRANSALDOLASE-LIKE PROTEIN (20 KD PHOSPHOPROTEIN ORFU) (CSJ9) [Bacillus subtilis]
876	Bt1Gc1205	Bt1G1245	65-1612	g4835822	1222	1146	2.80E-116	48	100	(AF102174) glycine betaine transporter BetL [Listeria monocytogenes]
877	Bt1Gc1204	Bt1G1246	1355-1	g2633180	77	135	3.80E-09	32	67	(Z99108) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
877	Bt1Gc1204	Bt1G1247	1656-1236	g2633808	408	379	5.20E-35	62	35	(Z99111) similar to hypothetical proteins [Bacillus subtilis]
878	Bt1Gc1206	Bt1G1248	1-308	g71129	358	156	2.20E-11	77	85	ribosomal protein L7/L12 - Bacillus subtilis []
878	Bt1Gc1206	Bt1G1249	388-984	g586914	678	545	1.30E-52	64	100	HYPOTHETICAL 22.5 KD PROTEIN IN RPLL-RPOB INTERGENIC REGION (P23) (ORF23) [Bacillus subtilis]
878	Bt1Gc1206	Bt1G1250	1279-1645	g585920	577	593	1.30E-56	92	10	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Bacillus subtilis]
879	Bt1Gc1207	Bt1G1251	575-1	g133949	849	846	1.70E-84	82	96	30S RIBOSOMAL PROTEIN S4 (BS4) [Bacillus subtilis]
879	Bt1Gc1207	Bt1G1252	963-2056	g1708975	804	860	5.60E-86	45	93	METHIONINE GAMMA-LYASE (L-METHIONINASE) [Pseudomonas putida]
880	Bt1Gc1209	Bt1G1253	1342-218	g1945051	511	409	3.50E-38	39	100	(U63928) L1 protein [Bacillus cereus]
881	Bt1Gc1208	Bt1G1254	1-2170	g1171128	1336	1416	6.80E-145	42	42	(U24657) saframycin Mx1 synthetase B [Myxococcus xanthus]
882	Bt1Gc1210	Bt1G1255	1-1008	g1934838	777	675	2.30E-66	43	74	(Z93940) unknown [Bacillus subtilis]
882	Bt1Gc1210	Bt1G1256	1852-1211	g551727	865	896	8.60E-90	80	70	(M57689) sporulation protein [Bacillus subtilis]
883	Bt1Gc1211	Bt1G1257	3163-681	g3821797	150	233	9.70E-19	35	35	(D10594) chitinase D precursor [Bacillus circulans]
884	Bt1Gc1212	Bt1G1258	959-39	g400209	834	688	9.40E-68	51	100	MEMBRANE-BOUND PROTEIN LYTR [Bacillus subtilis]
884	Bt1Gc1212	Bt1G1259	1782-1086	g1730193	1040	1042	2.90E-105	82	68	UDP-GLUCOSE 4-EPIMERASE (GALACTOWALDENASE) (UDP-GALACTOSE 4-EPIMERASE) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
885	Bt1Gc1215	Bt1G1260	1-351	g1176955	270	242	3.60E-20	43	29	HYPOTHETICAL 45.8 KD PROTEIN IN ACDA-NARI INTERGENIC REGION [Bacillus subtilis]
885	Bt1Gc1215	Bt1G1261	387-1493	g1176954	261	302	7.60E-27	59	32	HYPOTHETICAL 36.9 KD PROTEIN IN ACDA 5'REGION [Bacillus subtilis]
885	Bt1Gc1215	Bt1G1262	1493-787	g225559	639	676	1.80E-66	85	33	ORF IS231C [Bacillus thuringiensis]
886	Bt1Gc1216	Bt1G1263	881-222	g544315	423	346	1.60E-31	38	100	FLAGELLAR BIOSYNTHETIC PROTEIN FLIP [Bacillus subtilis]
886	Bt1Gc1216	Bt1G1264	2155-1333	g4322005	110	166	2.00E-12	26	99	(AF069392) polar flagellar switch protein [Vibrio parahaemolyticus]
886	Bt1Gc1216	Bt1G1265	2659-1645	g120344	261	362	3.30E-33	27	100	FLAGELLAR MOTOR SWITCH PROTEIN FLIM [Bacillus subtilis]
887	Bt1Gc1214	Bt1G1266	394-1	g1730930	175	238	4.60E-20	36	68	HYPOTHETICAL 23.7 KD PROTEIN IN ILVD-THYB INTERGENIC REGION [Bacillus subtilis]
887	Bt1Gc1214	Bt1G1267	2390-1875	g2293199	490	498	1.30E-47	53	72	(AF008220) transporter [Bacillus subtilis]
888	Bt1Gc1217	Bt1G1268	1161-278	g4104606	485	526	1.40E-50	37	75	(AF036967) putative histidine kinase [Lactobacillus sakei]
888	Bt1Gc1217	Bt1G1269	2083-1812	g4481749	200	223	1.80E-18	45	35	(AF007865) BacR [Bacillus licheniformis]
889	Bt1Gc1219	Bt1G1270	1455-1	g1169981	1362	1314	4.40E-134	54	94	GLUCONOKINASE (GLUCONATE KINASE) [Bacillus licheniformis]
890	Bt1Gc1222	Bt1G1271	69-1463	g1352421	795	831	6.60E-83	39	100	MENAQUINONE-SPECIFIC ISOCHORISMATE SYNTHASE [Bacillus subtilis]
891	Bt1Gc1224	Bt1G1272	1261-3034	g1731038	198	143	1.10E-09	48	40	HYPOTHETICAL 28.3 KD PROTEIN IN COMGG-SINR INTERGENIC REGION [Bacillus subtilis]
891	Bt1Gc1224	Bt1G1273	2251-1925	g134504	354	342	4.40E-31	67	99	SINR PROTEIN [Bacillus subtilis]
892	Bt1Gc1220	Bt1G1274	1-803	g1171678	973	920	2.50E-92	70	53	NADH DEHYDROGENASE SUBUNIT 5 (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5) [Bacillus subtilis]
892	Bt1Gc1220	Bt1G1275	1043-3313	g2632452	2706	2610	2.00E-271	69	100	(Z99104) ybcD [Bacillus subtilis]
893	Bt1Gc1223	Bt1G1276	1926-1	g3122874	398	501	6.20E-48	34	71	D-3-PHOSPHOGLYCERATE DEHYDROGENASE (PGDH) [Methanococcus jannaschii]
893	Bt1Gc1223	Bt1G1277	2205-1123	g3660126	951	964	5.30E-97	50	100	Phosphoserine Aminotransferase From Bacillus Circulans Subsp. Alkalophilus []
894	Bt1Gc1221	Bt1G1278	121-2103	g730002	1481	1259	2.90E-128	46	100	METHYL-ACCEPTING CHEMOTAXIS PROTEIN MCPA (H1) [Bacillus subtilis]
895	Bt1Gc1213	Bt1G1279	1607-957	g586859	782	785	5.00E-78	65	100	HYPOTHETICAL 25.4 KD PROTEIN IN SERS-DNAZ INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
895	Bt1Gc1213	Bt1G1280	2254-1634	g586860	554	555	1.20E-53	50	100	HYPOTHETICAL 24.1 KD PROTEIN IN SERS-DNAZ INTERGENIC REGION [Bacillus subtilis]
895	Bt1Gc1213	Bt1G1281	2925-2383	g586862	490	497	1.60E-47	53	99	HYPOTHETICAL 20.8 KD PROTEIN IN SERS-DNAZ INTERGENIC REGION [Bacillus subtilis]
896	Bt1Gc1225	Bt1G1282	310-68	g1731032	271	280	1.60E-24	63	99	HYPOTHETICAL 9.4 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
896	Bt1Gc1225	Bt1G1283	1337-2371	g1332095	1029	1042	2.90E-105	56	97	COMG OPERON PROTEIN 1 [Bacillus subtilis]
897	Bt1Gc1226	Bt1G1284	308-1	g225559	297	381	3.20E-35	80	22	ORF IS231C [Bacillus thuringiensis]
897	Bt1Gc1226	Bt1G1285	874-2079	g2293242	1651	1667	1.70E-171	78	100	(AF008220) arginine succinate synthase [Bacillus subtilis]
897	Bt1Gc1226	Bt1G1286	2079-2970	g2293243	1149	1086	6.30E-110	74	64	(AF008220) arginine succinate lyase [Bacillus subtilis]
898	Bt1Gc1230	Bt1G1287	247-587	g1175980	143	184	2.40E-14	43	64	HYPOTHETICAL PROTEIN HI0906 [Haemophilus influenzae Rd]
898	Bt1Gc1230	Bt1G1288	958-2850	g2632518	147	202	3.00E-16	35	39	(Z99105) similar to transcriptional regulator (AraC/XylS family) [Bacillus subtilis]
898	Bt1Gc1230	Bt1G1289	2888-3421	g1652543	108	158	4.00E-11	29	37	(D90906) hypothetical protein [Synecocystis sp.]
899	Bt1Gc1228	Bt1G1290	1455-1744	g2462116	219	263	1.00E-22	47	58	(Y11139) ORF2 [Bacillus cereus]
899	Bt1Gc1228	Bt1G1291	2155-2499	g2443243	287	256	5.70E-22	45	99	(D86417) YfIT [Bacillus subtilis]
900	Bt1Gc1232	Bt1G1292	1352-906	g121511	159	169	9.40E-13	31	99	COMG OPERON PROTEIN 4 PRECURSOR [Bacillus subtilis]
900	Bt1Gc1232	Bt1G1293	1630-1340	g121510	176	128	2.10E-08	43	99	COMG OPERON PROTEIN 3 PRECURSOR [Bacillus subtilis]
900	Bt1Gc1232	Bt1G1294	2490-1648	g116546	469	351	4.90E-32	32	87	COMG OPERON PROTEIN 2 [Bacillus subtilis]
901	Bt1Gc1227	Bt1G1295	951-54	g1730895	813	846	1.70E-84	53	97	POTENTIAL 5'-3' EXONUCLEASE [Bacillus subtilis]
901	Bt1Gc1227	Bt1G1296	1444-1098	g543864	147	249	5.30E-20	44	12	MG(2+) TRANSPORT ATPASE, P-TYPE 1 [Salmonella typhimurium]
902	Bt1Gc1231	Bt1G1297	2232-1538	g2634109	785	808	1.80E-80	63	100	(Z99113) ymaE [Bacillus subtilis]
903	Bt1Gc1233	Bt1G1298	2653-340	g1881371	1009	986	2.50E-99	36	70	(AB001488) PROBABLE TRANSPORT PROTEIN, SIMILAR PRODUCT IN MYCOBACTERIUM LEPRAE. [Bacillus subtilis]
904	Bt1Gc1238	Bt1G1299	1-2249	g2635792	84	227	4.20E-17	22	65	(Z99120) similar to oligoendopeptidase [Bacillus subtilis]
904	Bt1Gc1238	Bt1G1300	813-2065	g3323354	325	459	1.70E-43	25	71	(AF001270) oligoendopeptidase F, putative [Treponema pallidum]
905	Bt1Gc1234	Bt1G1301	416-1	g2266425	477	493	4.40E-47	67	31	(Y13917) yngH [Bacillus subtilis]
905	Bt1Gc1234	Bt1G1302	1577-438	g2266423	1481	1440	1.90E-147	74	100	(Y13917) yngJ [Bacillus subtilis]
905	Bt1Gc1234	Bt1G1303	2431-1815	g4456867	226	277	3.40E-24	27	65	(AJ224978) ORF 319 [Salmonella typhimurium]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
906	Bt1Gc1229	Bt1G1304	1-1010	g2769532	226	280	1.60E-24	42	61	(AJ223073) phosphate regulatory protein [Bradyrhizobium japonicum]
906	Bt1Gc1229	Bt1G1305	1-1778	g2182990	335	377	8.50E-35	32	53	(U81485) histidine kinase [Lactococcus lactis cremoris]
907	Bt1Gc1236	Bt1G1306	1-752	g2495393	358	439	2.30E-41	39	85	TRANSCRIPTIONAL REGULATOR SOXR [Arthrobacter sp.]
907	Bt1Gc1236	Bt1G1307	2219-1278	g585208	745	832	5.20E-83	52	100	GLYCEROPHOSPHORYL DIESTER PHOSPHODIESTERASE (GLYCEROPHOSPHODIESTER PHOSPHODIESTERASE) [Bacillus subtilis]
907	Bt1Gc1236	Bt1G1308	3440-2581	g769829	469	505	2.30E-48	37	51	(M80674) hydroxylase [Streptomyces glaucescens]
908	Bt1Gc1237	Bt1G1309	571-1116	g2633543	330	351	4.90E-32	38	99	(Z99110) similar to ribosomal-protein-alanine N-acetyltransferase [Bacillus subtilis]
908	Bt1Gc1237	Bt1G1310	1235-2173	g1731068	1248	1213	2.20E-123	78	84	HYPOTHETICAL 39.7 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
909	Bt1Gc1239	Bt1G1311	626-1	g4584151	1020	738	4.80E-73	98	92	(AJ010111) cytochrome caa3 oxidase subunit II [Bacillus cereus]
909	Bt1Gc1239	Bt1G1312	1643-723	g4584150	1603	1498	1.40E-153	100	100	(AJ010111) cytochrome caa3-oxidase assembly factor [Bacillus cereus]
909	Bt1Gc1239	Bt1G1313	2141-2874	g4584149	1234	1173	3.80E-119	98	78	(AJ010111) cytochrome aa3 controlling protein [Bacillus cereus]
910	Bt1Gc1235	Bt1G1314	3187-1	g2633724	562	624	5.70E-61	48	35	(Z99111) similar to two-component sensor histidine kinase [Bacillus subtilis]
910	Bt1Gc1235	Bt1G1315	3033-2173	g1894770	401	440	1.80E-41	36	100	(Z92954) product similar to Bacillus subtilis YxeH and YcsE proteins and to E. coli YidA protein [Bacillus subtilis]
911	Bt1Gc1241	Bt1G1316	1488-1	g2619051	1586	1489	1.20E-152	60	84	(AF027868) RecQ homolog [Bacillus subtilis]
912	Bt1Gc1240	Bt1G1317	2188-140	g1072970	618	668	1.20E-65	44	78	dhIR protein - Xanthobacter autotrophicus [Xanthobacter autotrophicus]
913	Bt1Gc1245	Bt1G1318	2490-1765	g1894750	744	759	2.80E-75	60	100	(Z92952) product similar to E.coli YjaF protein [Bacillus subtilis]
914	Bt1Gc1242	Bt1G1319	169-1004	g2226222	351	434	7.80E-41	34	100	(Y14082) hypothetical protein [Bacillus subtilis]
914	Bt1Gc1242	Bt1G1320	1880-1315	g418449	126	153	4.70E-11	24	99	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN HEMY-GLTT INTERGENIC REGION (ORFA) [Bacillus subtilis]
915	Bt1Gc1244	Bt1G1321	701-150	g729934	446	477	2.20E-45	49	99	SIGNAL PEPTIDASE I (SPASE I) (LEADER PEPTIDASE I) [Bacillus caldolyticus]
915	Bt1Gc1244	Bt1G1322	2017-1	g4680703	304	352	3.80E-32	34	88	(AF132966) CGI-32 protein [Homo sapiens]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvrq	NCBI gi description
915	Bt1Gc1244	Bt1G1323	2302-1625	g1881346	527	528	8.50E-51	45	100	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
916	Bt1Gc1246	Bt1G1324	19-966	g400209	516	543	2.20E-52	42	100	MEMBRANE-BOUND PROTEIN LYTR [Bacillus subtilis]
916	Bt1Gc1246	Bt1G1325	2416-2724	g2293270	243	244	1.10E-20	48	43	(AF008220) signal transduction regulator [Bacillus subtilis]
917	Bt1Gc1248	Bt1G1326	698-411	g113009	265	253	1.20E-21	49	99	TRANSCRIPTION STATE REGULATORY PROTEIN ABRB [Bacillus subtilis]
917	Bt1Gc1248	Bt1G1327	2096-2288	g4096799	64	112	1.00E-06	31	36	(U40158) orfx; function unknown; similar to response regulators of two-component regulatory systems; Method: conceptual translation supplied by author. [Staphylococcus carnosus]
918	Bt1Gc1249	Bt1G1328	1025-1	g2496635	382	471	7.60E-44	41	26	HYPOTHETICAL 102.8 KD PROTEIN Y4GI [Rhizobium sp. NGR234]
918	Bt1Gc1249	Bt1G1329	1947-1105	g1805396	244	275	5.50E-24	27	100	(D50453) yegQ [Bacillus subtilis]
918	Bt1Gc1249	Bt1G1330	2542-2008	g1805397	329	373	2.30E-34	40	62	(D50453) homologue of unidentified protein of E. coli [Bacillus subtilis]
919	Bt1Gc1247	Bt1G1331	172-1981	g4151933	653	795	4.30E-79	34	100	(AF110737) Rhsc [Sinorhizobium meliloti]
920	Bt1Gc1251	Bt1G1332	1799-1	g2633600	482	485	3.10E-46	56	54	(Z99110) N-acetylmuramoyl-L-alanine amidase [Bacillus subtilis]
920	Bt1Gc1251	Bt1G1333	2062-1790	g4126631	129	128	2.10E-08	30	67	(AB016282) ORF45 [bacteriophage phi-105]
921	Bt1Gc1254	Bt1G1334	1-1020	g2649875	54	126	3.40E-08	28	72	(AE001054) ribosomal protein S18 alanine acetyltransferase [Archaeoglobus fulgidus]
921	Bt1Gc1254	Bt1G1335	2440-1284	g3913040	771	850	6.40E-85	43	99	PUTATIVE ALANINE RACEMASE [Bacillus subtilis]
922	Bt1Gc1253	Bt1G1336	1-223	g2738159	153	171	5.80E-13	46	28	(U91841) MotB homolog [Bacillus firmus]
922	Bt1Gc1253	Bt1G1337	351-710	g620085	441	454	5.90E-43	74	99	(X76170) cheY [Listeria monocytogenes]
922	Bt1Gc1253	Bt1G1338	846-2229	g2500757	1130	1124	5.90E-114	59	67	CHEMOTAXIS PROTEIN CHEA [Listeria monocytogenes]
923	Bt1Gc1256	Bt1G1339	507-2318	g1296829	1041	1149	1.30E-116	42	100	(X89922) maturase-related protein [Lactococcus lactis]
923	Bt1Gc1256	Bt1G1340	2827-3336	g1750115	769	772	1.20E-76	82	99	(U66480) YnaD [Bacillus subtilis]
924	Bt1Gc1257	Bt1G1341	400-1	g2501608	362	377	8.50E-35	51	78	HYPOTHETICAL 19.2 KD PROTEIN IN RPH-ILVB INTERGENIC REGION [Bacillus subtilis]
924	Bt1Gc1257	Bt1G1342	1025-426	g1770060	547	499	1.00E-47	57	99	(Z75208) hypothetical protein [Bacillus subtilis]
924	Bt1Gc1257	Bt1G1343	1420-1031	g322197	379	406	7.20E-38	62	53	phosphate-dependent exoribonuclease - Bacillus subtilis []
925	Bt1Gc1258	Bt1G1344	2021-1308	g1176948	536	549	5.10E-53	43	100	HYPOTHETICAL 27.6 KD PROTEIN IN FNR-NARG INTERGENIC REGION [Bacillus subtilis]
926	Bt1Gc1259	Bt1G1345	1871-671	g1731041	119	181	5.00E-14	34	99	HYPOTHETICAL 14.6 KD PROTEIN IN GCVT-



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
926	Bt1Gc1259	Bt1G1346	2085-1297	g1707680	365	497	1.60E-47	38	100	SPOIIIAA INTERGENIC REGION [Bacillus subtilis]
926	Bt1Gc1259	Bt1G1347	3345-2111	g1652676	971	1042	2.90E-105	47	100	(Y08256) orf c06001 [Sulfolobus solfataricus]
927	Bt1Gc1261	Bt1G1348	1-2428	g128494	72	180	1.20E-12	22	29	(D90907) hypothetical protein [Synechocystis sp.]
927	Bt1Gc1261	Bt1G1349	1-2428	g2633696	747	746	6.70E-74	36	71	NODULATION PROTEIN V [Bradyrhizobium japonicum]
928	Bt1Gc1252	Bt1G1350	534-1160	g2226262	316	361	4.20E-33	39	100	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
928	Bt1Gc1252	Bt1G1351	1480-2101	g2226177	536	559	4.40E-54	49	100	(Y14084) competence transcription factor [Bacillus subtilis]
929	Bt1Gc1260	Bt1G1352	340-2256	g135176	2316	2231	2.90E-231	66	100	(Y14081) hypothetical protein [Bacillus subtilis]
929	Bt1Gc1260	Bt1G1353	2515-3146	g2293177	564	582	1.60E-56	51	83	THREONYL-TRNA SYNTHETASE 2 (THREONINE--TRNA LIGASE) (THRRS) [Bacillus subtilis]
930	Bt1Gc1263	Bt1G1354	1121-297	g2499116	335	371	3.70E-34	29	100	(AF008220) transporter [Bacillus subtilis]
930	Bt1Gc1263	Bt1G1355	2867-2104	g143324	998	998	1.30E-100	74	66	VANCOMYCIN B-TYPE RESISTANCE PROTEIN
931	Bt1Gc1264	Bt1G1356	1-905	g887872	203	355	1.80E-32	35	82	VANW [Enterococcus faecalis]
931	Bt1Gc1264	Bt1G1357	2171-1447	g116300	339	423	4.30E-39	40	34	(M37169) APase I [Bacillus licheniformis]
931	Bt1Gc1264	Bt1G1358	2781-2179	g3308998	465	516	1.60E-49	51	99	(U25453) phosphotidyl inositol-specific phospholipase C [Listeria monocytogenes]
932	Bt1Gc1262	Bt1G1359	1-399	g3183496	239	253	1.20E-21	35	65	CHITINASE A1 PRECURSOR [Bacillus circulans]
932	Bt1Gc1262	Bt1G1360	1936-470	g1170977	1464	1378	7.20E-141	57	100	(AB015998) CBP21 precursor [Serratia marcescens]
932	Bt1Gc1262	Bt1G1361	2865-1972	g140739	407	477	2.20E-45	35	95	HYPOTHETICAL 23.4 KD PROTEIN IN AAPA-SIGV INTERGENIC REGION [Bacillus subtilis]
932	Bt1Gc1262	Bt1G1362	3925-2873	g2635778	641	560	3.50E-54	43	63	PROBABLE METHYLMALONATE-SEMIALDEHYDE DEHYDROGENASE [Bacillus subtilis]
933	Bt1Gc1266	Bt1G1363	1128-1520	g2633504	571	571	2.40E-55	81	99	HYPOTHETICAL 31.0 KD PROTEIN IN RNPB-SOHA INTERGENIC REGION (ORF 2) [Escherichia coli]
933	Bt1Gc1266	Bt1G1364	2247-1591	g2633505	550	396	8.30E-37	49	100	(Z99120) similar to butyryl-CoA dehydrogenase [Bacillus subtilis]
934	Bt1Gc1267	Bt1G1366	750-1	g1170997	361	166	2.40E-12	33	93	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
934	Bt1Gc1267	Bt1G1367	1780-1598	g4894352	159	164	3.20E-12	52	98	HYPOTHETICAL 29.3 KD PROTEIN IN CCPA 3'REGION (ORF1) [Bacillus megaterium]
934	Bt1Gc1267	Bt1G1368	3057-3401	g2467222	82	113	6.40E-06	25	27	(AF065404) pXO1-137 [Bacillus anthracis]
935	Bt1Gc1269	Bt1G1369	1009-107	g1168646	1210	1144	4.50E-116	74	100	(X92946) macrolide efflux protein [Lactococcus lactis]
935	Bt1Gc1269	Bt1G1370	1429-830	g1397290	248	320	9.40E-29	42	99	GTP-BINDING PROTEIN ERA HOMOLOG (BEX PROTEIN) [Bacillus subtilis]
										(U61949) Similar to cytidine deaminase. [Caenorhabditis elegans]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
935	Bt1Gc1269	Bt1G1371	1847-1506	g1708628	276	202	3.00E-16	49	99	DIACYLGLYCEROL KINASE (DAGK) (DIGLYCERIDE KINASE) (DGK) [Bacillus subtilis]
935	Bt1Gc1269	Bt1G1372	2320-1850	g1731000	528	367	9.80E-34	64	99	HYPOTHETICAL 17.8 KD PROTEIN IN PHOH-DGKA INTERGENIC REGION [Bacillus subtilis]
935	Bt1Gc1269	Bt1G1373	2443-2326	g1730999	117	134	7.60E-08	63	5	HYPOTHETICAL 79.2 KD PROTEIN IN PHOH-DGKA INTERGENIC REGION [Bacillus subtilis]
936	Bt1Gc1270	Bt1G1374	482-1	g729356	394	366	1.20E-33	45	69	DNA REPLICATION PROTEIN DNAD [Bacillus subtilis]
936	Bt1Gc1270	Bt1G1375	1787-612	g2492839	1522	1507	1.50E-154	77	100	ASPARTATE AMINOTRANSFERASE (TRANSAMINASE A) (ASPAT) [Bacillus stearothermophilus]
936	Bt1Gc1270	Bt1G1376	2315-1815	g1730935	188	210	4.20E-17	38	99	HYPOTHETICAL 17.9 KD PROTEIN IN DING-ASPB INTERGENIC REGION [Bacillus subtilis]
936	Bt1Gc1270	Bt1G1377	2487-2320	g1730934	203	203	2.30E-16	70	98	HYPOTHETICAL 6.6 KD PROTEIN IN DING-ASPB INTERGENIC REGION [Bacillus subtilis]
937	Bt1Gc1272	Bt1G1378	462-1	g2633171	430	448	2.60E-42	53	52	(Z99108) similar to hypothetical proteins [Bacillus subtilis]
937	Bt1Gc1272	Bt1G1379	2558-2034	g2633543	160	212	2.60E-17	28	99	(Z99110) similar to ribosomal-protein-alanine N-acetyltransferase [Bacillus subtilis]
937	Bt1Gc1272	Bt1G1380	2564-2034	g141186	240	259	2.70E-22	33	99	HYPOTHETICAL P20 PROTEIN [Bacillus licheniformis]
937	Bt1Gc1272	Bt1G1381	2948-2720	g462484	131	169	9.30E-12	43	14	LONG-CHAIN-FATTY-ACID--COA LIGASE (LONG-CHAIN ACYL-COA SYNTHETASE) [Escherichia coli]
938	Bt1Gc1268	Bt1G1382	490-1	g1731048	378	397	6.50E-37	49	47	PUTATIVE PEPTIDASE IN GCVT-SPOIIIAA INTERGENIC REGION [Bacillus subtilis]
938	Bt1Gc1268	Bt1G1383	2556-841	g2462963	1851	1601	1.70E-164	62	100	(AJ000974) putative fibronectin-binding protein [Bacillus subtilis]
938	Bt1Gc1268	Bt1G1384	2656-3247	g2337795	581	568	1.80E-54	60	22	(Y13937) putative PaclL protein [Bacillus subtilis]
939	Bt1Gc1273	Bt1G1385	24-1874	g1339850	2621	2479	1.50E-257	79	100	(D85547) cytochrome c oxidase subunit I [Bacillus sp.]
940	Bt1Gc1276	Bt1G1386	737-63	g1934806	516	395	1.10E-36	48	100	(Z93936) unknown [Bacillus subtilis]
940	Bt1Gc1276	Bt1G1387	1630-851	g1730928	334	382	2.50E-35	30	100	HYPOTHETICAL 30.6 KD PROTEIN IN QCRC-DAPB INTERGENIC REGION PRECURSOR [Bacillus subtilis]
941	Bt1Gc1278	Bt1G1388	1144-1	g461914	317	400	3.10E-37	33	81	PENICILLIN-BINDING PROTEIN 5* PRECURSOR (D-ALANYL-D-ALANINE CARBOXYPEPTIDASE) (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) (PBP-5*) [Bacillus subtilis]
942	Bt1Gc1277	Bt1G1389	1-398	g3183185	517	488	1.50E-46	77	22	GTP-BINDING PROTEIN TYPA/BIPA HOMOLOG [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Prob	BlastP-Prob	% Ident	% Cvr	NCBI gi description
942	Bt1Gc1277	Bt1G1390	547-861	g2224767	312	319	1.20E-28	57	99	(Z97025) ylaH [Bacillus subtilis]
942	Bt1Gc1277	Bt1G1391	1388-1184	g2224768	184	160	8.40E-12	52	99	(Z97025) ylaI [Bacillus subtilis]
942	Bt1Gc1277	Bt1G1392	2233-2660	g2224770	574	590	2.30E-57	83	32	(Z97025) product similar to E. coli PhoH protein [Bacillus subtilis]
943	Bt1Gc1271	Bt1G1393	301-1	g98442	449	465	4.00E-44	86	46	recM protein - Bacillus subtilis []
943	Bt1Gc1271	Bt1G1394	642-322	g141372	443	353	3.00E-32	80	99	HYPOTHETICAL 11.8 KD PROTEIN IN DNAZ-RECR INTERGENIC REGION [Bacillus subtilis]
943	Bt1Gc1271	Bt1G1395	2356-671	g98292	1746	1699	6.90E-175	60	100	DNA-directed DNA polymerase (EC 2.7.7.7) III chain dnaX - Bacillus subtilis [Bacillus subtilis]
944	Bt1Gc1279	Bt1G1396	1379-88	g2507344	1575	1638	2.00E-168	77	36	DNA-DIRECTED RNA POLYMERASE BETA' CHAIN (TRANSCRIPTASE BETA' CHAIN) (RNA POLYMERASE BETA' SUBUNIT) []
945	Bt1Gc1274	Bt1G1397	2443-1528	g4583559	258	263	1.00E-22	24	100	(AJ005255) OxyR [Erwinia chrysanthemi]
945	Bt1Gc1274	Bt1G1398	2577-3092	g3256407	350	441	1.40E-41	49	45	(AP000001) 373aa long hypothetical sugar-binding transport ATP-binding protein [Pyrococcus horikoshii]
946	Bt1Gc1280	Bt1G1399	1854-210	g2828494	1853	1732	2.20E-178	66	100	NUCLEOTIDE BINDING PROTEIN EXPZ [Bacillus subtilis]
946	Bt1Gc1280	Bt1G1400	2977-2306	g1176994	516	563	1.70E-54	46	70	HYPOTHETICAL SENSOR-LIKE HISTIDINE KINASE IN IDH 3 REGION [Bacillus subtilis]
947	Bt1Gc1275	Bt1G1401	789-46	g113739	1086	1086	6.30E-110	80	100	METHIONINE AMINOPEPTIDASE (MAP) (PEPTIDASE M) [Bacillus subtilis]
947	Bt1Gc1275	Bt1G1402	1439-789	g125158	866	870	4.90E-87	75	100	ADENYLATE KINASE (ATP-AMP TRANSPHOSPHORYLASE) [Bacillus stearothermophilus]
947	Bt1Gc1275	Bt1G1403	1690-1502	g134408	260	261	1.70E-22	81	63	PREPROTEIN TRANSLOCASE SECY SUBUNIT [Bacillus stearothermophilus]
948	Bt1Gc1282	Bt1G1404	417-2277	g2633134	1483	1578	4.60E-162	50	100	(Z99108) transcriptional regulator [Bacillus subtilis]
949	Bt1Gc1281	Bt1G1405	150-1565	g1709059	1991	1942	1.20E-200	78	100	MMGE PROTEIN [Bacillus subtilis]
949	Bt1Gc1281	Bt1G1406	1587-2489	g1731059	1139	1092	1.50E-110	73	100	HYPOTHETICAL 33.1 KD PROTEIN IN MMGE-BFMBA INTERGENIC REGION [Bacillus subtilis]
949	Bt1Gc1281	Bt1G1407	2642-3186	g2635778	446	480	1.00E-45	48	30	(Z99120) similar to butyryl-CoA dehydrogenase [Bacillus subtilis]
950	Bt1Gc1283	Bt1G1408	121-2539	g1277135	374	374	1.80E-34	41	42	(U50978) kanamycin/gentamycin-resistance protein [Cloning vector pFW13]
950	Bt1Gc1283	Bt1G1409	686-1675	g1934645	179	215	1.30E-17	32	99	(U93876) hypothetical protein YrdC [Bacillus subtilis]
951	Bt1Gc1285	Bt1G1410	2294-2094	g1787753	220	221	4.00E-18	66	19	(AF000245) alcohol dehydrogenase [Escherichia coli]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
952	Bt1Gc1287	Bt1G1411	281-772	g22226150	532	497	1.60E-47	61	99	(Y14080) hypothetical protein [Bacillus subtilis]
952	Bt1Gc1287	Bt1G1412	771-1994	g22226151	1624	1222	2.40E-124	79	100	(Y14080) hypothetical protein [Bacillus subtilis]
953	Bt1Gc1284	Bt1G1413	1173-121	g1881258	921	598	3.30E-58	50	100	(AB001488) FUNCTION UNKNOWN, SIMILAR PRODUCTS IN SYNECHOCYSTIS AND H. INFLUENZAE. [Bacillus subtilis]
953	Bt1Gc1284	Bt1G1414	2187-1494	g1770004	881	609	2.20E-59	76	100	(Z75208) hypothetical protein [Bacillus subtilis]
954	Bt1Gc1288	Bt1G1415	669-1	g1894745	358	428	3.40E-40	39	80	(Z92952) unknown [Bacillus subtilis]
954	Bt1Gc1288	Bt1G1416	3313-1189	g1778501	132	242	3.70E-20	28	58	(U82598) enterochelin esterase [Escherichia coli]
955	Bt1Gc1290	Bt1G1417	1203-2724	g3334472	1317	1335	2.60E-136	51	66	5-METHYLTETRAHYDROPTEROYLTRIGLUTAMATE-HOMOCYSTEINE METHYLTRANSFERASE (METHIONINE SYNTHASE, VITAMIN-B12 INDEPENDENT ISOZYME) (COBALAMIN-INDEPENDENT METHIONINE SYNTHASE) (SUPEROXIDE-INDUCIBLE PROTEIN 9) (SOI9) [Bacillus subtilis]
956	Bt1Gc1291	Bt1G1418	325-1125	g1945654	362	402	1.90E-37	33	100	(Z94043) hypothetical protein [Bacillus subtilis]
956	Bt1Gc1291	Bt1G1419	3083-3367	g1575061	123	157	1.80E-11	34	41	(U57060) ScdA [Staphylococcus aureus]
957	Bt1Gc1292	Bt1G1420	300-1	g4894302	231	273	9.00E-24	56	26	(AF065404) pXO1-86 [Bacillus anthracis]
957	Bt1Gc1292	Bt1G1421	862-2345	g1730891	756	815	8.70E-81	35	41	HYPOTHETICAL 137.4 KD PROTEIN IN BC5A-DEGR INTERGENIC REGION [Bacillus subtilis]
958	Bt1Gc1289	Bt1G1422	1343-1	g1723606	806	616	4.00E-60	39	78	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN 1 IN GLVBC 3'REGION [Bacillus subtilis]
958	Bt1Gc1289	Bt1G1423	2168-1551	g731965	123	219	4.70E-18	30	99	HYPOTHETICAL 22.0 KD PROTEIN IN HXT11-HXT8 INTERGENIC REGION [Saccharomyces cerevisiae]
958	Bt1Gc1289	Bt1G1424	3092-2208	g2126561	282	356	1.40E-32	28	100	hypothetical protein 2 (downstream of hydrogenase cluster) - Anabaena variabilis [Anabaena variabilis]
958	Bt1Gc1289	Bt1G1425	3699-3373	g2443233	241	285	4.80E-25	51	48	(D86417) YfIK [Bacillus subtilis]
959	Bt1Gc1294	Bt1G1426	1424-1741	g3258025	118	112	1.90E-06	32	48	(AP000006) 215aa long hypothetical protein [Pyrococcus horikoshii]
959	Bt1Gc1294	Bt1G1427	2964-3704	g3258025	144	167	1.50E-12	29	100	(AP000006) 215aa long hypothetical protein [Pyrococcus horikoshii]
960	Bt1Gc1295	Bt1G1428	1-2553	g1894748	335	388	1.30E-35	45	29	(Z92952) product similar to Bacillus subtilis YxiD protein [Bacillus subtilis]
960	Bt1Gc1295	Bt1G1429	2549-2858	g1894749	87	129	1.60E-08	31	65	(Z92952) unknown [Bacillus subtilis]
961	Bt1Gc1297	Bt1G1430	622-1162	g1171921	316	388	5.80E-36	38	62	GLYCINE BETAIN-BINDING PROTEIN PRECURSOR

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
961	Bt1Gc1297	Bt1G1431	622-2491	g1171921	319	393	1.70E-36	38	65	[Bacillus subtilis] GLYCINE BETABINE-BINDING PROTEIN PRECURSOR
961	Bt1Gc1297	Bt1G1432	1796-1205	g3287962	637	583	1.30E-56	64	99	[Bacillus subtilis] ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT (ENDOPEPTIDASE CLP) (CASEINOLYTIC PROTEASE) (PROTEASE TD) [Bacillus subtilis]
961	Bt1Gc1297	Bt1G1433	1900-3641	g4262236	382	381	3.20E-35	39	100	(AC006200) putative ribose 5-phosphate isomerase [Arabidopsis thaliana]
961	Bt1Gc1297	Bt1G1434	3749-4189	g586830	433	447	3.30E-42	56	99	HYPOTHETICAL 16.6 KD PROTEIN IN COTF-TETB INTERGENIC REGION [Bacillus subtilis]
962	Bt1Gc1298	Bt1G1435	1268-42	g1176567	1601	1601	1.70E-164	74	100	HYPOTHETICAL PROCESSING PROTEASE (ORFP) [Bacillus subtilis]
962	Bt1Gc1298	Bt1G1436	2294-1355	g2634042	839	870	4.90E-87	52	98	(Z99112) alternate gene name: ymxI; similar to deacetylase [Bacillus subtilis]
963	Bt1Gc1299	Bt1G1437	1541-1	g2983239	296	348	7.70E-31	53	16	(AE000699) penicillin binding protein 1A [Aquifex aeolicus]
963	Bt1Gc1299	Bt1G1438	1414-28	g4582371	100	148	1.60E-10	31	89	(AL049573) MurT-like protein [Streptomyces coelicolor]
963	Bt1Gc1299	Bt1G1439	2331-1615	g2632998	1040	976	2.90E-98	85	100	(Z99107) similar to hypothetical proteins [Bacillus subtilis]
964	Bt1Gc1296	Bt1G1440	819-1	g586024	1096	849	6.70E-106	79	57	STAGE V SPORULATION PROTEIN R [Bacillus subtilis]
964	Bt1Gc1296	Bt1G1441	1512-943	g2633172	277	310	1.10E-27	32	99	(Z99108) similar to hypothetical proteins [Bacillus subtilis]
965	Bt1Gc1301	Bt1G1442	1222-1	g2293322	859	756	5.90E-75	44	93	(AF008220) branch-chain amino acid transporter [Bacillus subtilis]
965	Bt1Gc1301	Bt1G1443	2009-1443	g2833396	890	890	3.70E-89	89	99	COTJC PROTEIN [Bacillus subtilis]
965	Bt1Gc1301	Bt1G1444	2331-2025	g2833395	263	205	1.40E-16	52	99	COTJB PROTEIN [Bacillus subtilis]
965	Bt1Gc1301	Bt1G1445	2526-2278	g2833394	239	276	4.30E-24	52	99	COTJA PROTEIN [Bacillus subtilis]
966	Bt1Gc1293	Bt1G1446	3385-2391	g1637534	398	438	2.90E-41	31	100	(U73857) cyn operon transcriptional activator [Escherichia coli]
966	Bt1Gc1293	Bt1G1447	3457-3670	g2443233	169	188	9.10E-15	49	32	(D86417) YfK [Bacillus subtilis]
967	Bt1Gc1303	Bt1G1448	2268-1862	g119547	156	232	6.20E-19	30	32	ERYTHROMYCIN ESTERASE TYPE II [Escherichia coli]
968	Bt1Gc1305	Bt1G1449	1284-1	g732119	131	270	1.90E-23	27	88	HYPOTHETICAL 39.8 KD PROTEIN IN OSMY-DEOC INTERGENIC REGION (O357) [Escherichia coli]
968	Bt1Gc1305	Bt1G1450	1482-983	g2828809	211	269	2.40E-23	32	99	(AF002191) YhzA homolog [Bacillus subtilis]
969	Bt1Gc1304	Bt1G1451	1-468	g1171068	599	608	2.80E-59	74	36	PROBABLE UDP-N-ACETYLGLUCOSAMINE 1- CARBOXYVINYLTANSFERASE (ENOYLPIRUVATE TRANSFERASE) (UDP-N-ACETYLGLUCOSAMINE

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
969	Bt1Gc1304	Bt1G1452	588-1553	g418587	1414	1427	4.60E-146	86	100	ENOLPYRUVYL TRANSFERASE (EPT) [Bacillus subtilis]
969	Bt1Gc1304	Bt1G1453	1823-3102	g143434	1802	1825	3.10E-188	84	100	HYPOTHETICAL 34.0 KD PROTEIN IN RHO-MURA INTERGENIC REGION (ORFQ) [Bacillus subtilis]
970	Bt1Gc1306	Bt1G1454	930-48	g2293237	356	365	1.60E-33	50	99	(M97678) Rho Factor [Bacillus subtilis]
970	Bt1Gc1306	Bt1G1455	1139-1519	g1731010	103	99	2.50E-05	31	99	(AF008220) YnfJ [Bacillus subtilis]
970	Bt1Gc1306	Bt1G1456	2923-1589	g2635594	1350	1307	2.40E-133	55	100	HYPOTHETICAL 13.9 KD PROTEIN IN COCA-SODA INTERGENIC REGION [Bacillus subtilis]
971	Bt1Gc1307	Bt1G1457	2523-664	g2633811	1841	1557	7.70E-160	60	100	(Z99119) similar to Na <sup>+</sup> -transporting ATP synthase [Bacillus subtilis]
971	Bt1Gc1307	Bt1G1458	2711-2540	g2633810	110	126	1.30E-07	47	19	(Z99111) phosphotransferase system (PTS) fructose-specific enzyme IIBC component [Bacillus subtilis]
972	Bt1Gc1308	Bt1G1459	1-1093	g2634337	1017	912	1.70E-91	54	81	(Z99111) fructose-1-phosphate kinase [Bacillus subtilis]
972	Bt1Gc1308	Bt1G1460	2185-3107	g3080755	149	193	3.10E-15	25	88	(AF016483) APH(2'')-Id [Enterococcus casseliflavus]
973	Bt1Gc1309	Bt1G1461	263-1	g2829805	167	186	1.50E-14	41	32	HYPOTHETICAL 30.5 KD PROTEIN IN GABP-GUAA INTERGENIC REGION [Bacillus subtilis]
973	Bt1Gc1309	Bt1G1462	1606-647	g1708641	1018	1035	1.60E-104	62	100	(U51115) YeaC [Bacillus subtilis]
974	Bt1Gc1310	Bt1G1463	1612-83	g1075130	302	144	1.00E-15	27	100	periplasmic-binding-protein-dependent iron transport protein (sfuB) homolog - Haemophilus influenzae (strain Rd KW20) [Haemophilus influenzae Rd]
974	Bt1Gc1310	Bt1G1464	1736-910	g2766195	239	211	3.30E-17	25	100	(U75349) putative permease BhiE [Brachyspira hyodysenteriae]
974	Bt1Gc1310	Bt1G1465	2251-1745	g4337125	99	148	1.60E-10	26	99	(U75349) periplasmic-iron-binding protein BhiA [Brachyspira hyodysenteriae]
974	Bt1Gc1310	Bt1G1466	2752-706	g4894577	86	159	4.20E-11	27	52	(AF117273) iron-binding protein precursor [Ehrlichia chaffeensis]
974	Bt1Gc1310	Bt1G1467	3248-1609	g2766194	175	231	4.50E-19	44	28	(U75349) putative ABC transporter BhiD [Brachyspira hyodysenteriae]
975	Bt1Gc1311	Bt1G1468	1-1361	g2829802	599	380	2.70E-65	34	100	HYPOTHETICAL 53.3 KD PROTEIN IN SFP-GERKA INTERGENIC REGION [Bacillus subtilis]
975	Bt1Gc1311	Bt1G1469	2060-1767	g4894353	218	194	2.10E-15	48	99	(AF065404) pXO1-138 [Bacillus anthracis]
976	Bt1Gc1314	Bt1G1470	1-1563	g3323079	518	258	1.40E-19	30	90	(AE001248) conserved hypothetical integral membrane protein [Treponema pallidum]
976	Bt1Gc1314	Bt1G1471	2268-1634	g2226252	471	458	2.20E-43	46	64	(Y14084) hypothetical protein [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
977	Bt1Gc1313	Bt1G1472	1916-387	g2633324	551	594	8.60E-58	45	100	(Z99109) similar to 3-hydroxybutyryl-CoA dehydratase [Bacillus subtilis]
978	Bt1Gc1316	Bt1G1473	214-546	g1176702	462	462	8.40E-44	77	99	HYPOTHETICAL 13.0 KD PROTEIN IN QCRC-DAPB INTERGENIC REGION [Bacillus subtilis]
978	Bt1Gc1316	Bt1G1474	549-1352	g1706300	901	875	1.40E-87	65	100	DIHYDRODIPICOLINATE REDUCTASE [Bacillus subtilis]
978	Bt1Gc1316	Bt1G1475	1370-1783	g1176703	475	496	2.10E-47	68	99	METHYLGLYOXAL SYNTHASE [Bacillus subtilis]
978	Bt1Gc1316	Bt1G1476	1777-2442	g1176704	668	684	2.50E-67	61	100	HYPOTHETICAL 24.8 KD PROTEIN IN DAPB-PAPS INTERGENIC REGION [Bacillus subtilis]
978	Bt1Gc1316	Bt1G1477	2475-2604	g1730929	119	134	2.80E-08	60	11	HYPOTHETICAL 42.0 KD PROTEIN IN DAPB-PAPS INTERGENIC REGION [Bacillus subtilis]
979	Bt1Gc1302	Bt1G1478	1-573	g2612898	516	522	3.70E-50	57	59	(AF015825) YisP-like protein [Bacillus subtilis]
979	Bt1Gc1302	Bt1G1479	614-3114	g1001551	57	202	1.60E-15	24	57	(D64000) hypothetical protein [Synecocystis sp.]
980	Bt1Gc1315	Bt1G1480	834-1815	g1703442	779	844	2.80E-84	51	100	ASPARTATE--AMMONIA LIGASE (ASPARAGINE SYNTHETASE A) [Lactobacillus delbrueckii]
980	Bt1Gc1315	Bt1G1481	2345-1881	g1749770	308	351	4.90E-32	46	52	(Y09946) transposase [Bacillus thuringiensis]
981	Bt1Gc1312	Bt1G1482	382-1809	g2982907	1090	1055	1.20E-106	46	100	(AE000677) aldehyde dehydrogenase [Aquifex aeolicus]
981	Bt1Gc1312	Bt1G1483	2304-1909	g2635672	267	273	9.00E-24	39	99	(Z99120) yueI [Bacillus subtilis]
981	Bt1Gc1312	Bt1G1484	2555-2614	g2127280	108	108	1.80E-05	100	5	histidine protein kinase (Tn5401) - Bacillus thuringiensis [Bacillus thuringiensis]
982	Bt1Gc1318	Bt1G1485	397-1	g3258251	174	201	3.80E-16	34	48	(AP000007) 284aa long hypothetical cobalt transport ATP-binding protein [Pyrococcus horikoshii]
982	Bt1Gc1318	Bt1G1486	2326-532	g2632414	85	145	6.70E-10	30	46	(Z99104) ybaF [Bacillus subtilis]
983	Bt1Gc1317	Bt1G1488	2455-2102	g2497392	584	585	7.80E-57	96	47	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
984	Bt1Gc1319	Bt1G1489	1-493	g1204036	201	217	7.70E-18	29	58	(L40490) urease [Ureaplasma urealyticum]
984	Bt1Gc1319	Bt1G1490	515-1135	g2498126	257	186	1.50E-14	34	100	PUTATIVE TRANSPORTER PROTEIN AMIS2 [Rhodococcus sp. R312]
984	Bt1Gc1319	Bt1G1491	1149-2097	g4154878	504	277	6.80E-41	40	100	(AE001470) HIGH-AFFINITY NICKEL-TRANSPORT PROTEIN [Helicobacter pylori J99]
984	Bt1Gc1319	Bt1G1492	3766-4160	g141186	187	180	6.40E-14	36	74	HYPOTHETICAL P20 PROTEIN [Bacillus licheniformis]
985	Bt1Gc1322	Bt1G1493	236-679	g1303698	255	278	2.60E-24	43	99	(D84432) Btd [Bacillus subtilis]
985	Bt1Gc1322	Bt1G1494	2601-1080	g953179	376	445	5.30E-42	37	100	(Z37980) ORF14 [Escherichia coli]
986	Bt1Gc1325	Bt1G1495	2757-2050	g399406	246	212	2.60E-17	25	85	DAUNORUBICIN RESISTANCE TRANSMEMBRANE PROTEIN [Streptomyces peucetius]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
987	Bt1Gc1323	Bt1G1496	1687-1	g1708794	2457	2314	4.70E-240	85	92	GTP-BINDING PROTEIN LEPA [Bacillus subtilis]
988	Bt1Gc1326	Bt1G1497	2119-287	g2196513	169	358	8.80E-33	35	70	(U77778) putative membrane protein [Staphylococcus epidermidis]
988	Bt1Gc1326	Bt1G1498	1447-2142	g2633027	293	297	2.60E-26	35	100	(Z99107) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
988	Bt1Gc1326	Bt1G1499	3526-2140	g2314344	620	493	4.40E-47	32	100	(AE000624) conserved hypothetical integral membrane protein [Helicobacter pylori 26695]
988	Bt1Gc1326	Bt1G1500	3536-2177	g2688416	175	182	2.60E-11	21	100	(AE001152) conserved hypothetical integral membrane protein [Borrelia burgdorferi]
988	Bt1Gc1326	Bt1G1501	4356-3680	g1881353	262	309	1.40E-27	29	83	(AB001488) PROBABLE REGULATORY PROTEIN, SIMILAR TO MULTIDRUG-EFFLUX TRANSPORTER REGULATORS IN B.SUBTILIS. [Bacillus subtilis]
989	Bt1Gc1321	Bt1G1502	2698-1	g3025120	969	925	7.30E-93	37	93	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YDIF [Bacillus subtilis]
990	Bt1Gc1324	Bt1G1503	342-1	g2633840	314	347	1.30E-31	58	37	(Z99111) ykd [Bacillus subtilis]
990	Bt1Gc1324	Bt1G1504	3122-2445	g2649101	444	423	1.10E-39	40	100	(AE001001) ABC transporter, ATP-binding protein [Archaeoglobus fulgidus]
991	Bt1Gc1327	Bt1G1505	1656-1	g1321625	2084	2064	1.50E-213	67	99	(D84648) exo-alpha-1,4-glucosidase [Bacillus stearothermophilus]
991	Bt1Gc1327	Bt1G1506	1795-2273	g1673402	214	304	9.30E-27	42	29	(Z82044) hypothetical 54.4 kd protein [Bacillus subtilis]
991	Bt1Gc1327	Bt1G1507	2273-2177	g2497392	148	164	3.20E-12	97	13	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
992	Bt1Gc1330	Bt1G1508	1-1319	g1934619	264	349	7.90E-32	32	93	(U93874) YrhO [Bacillus subtilis]
992	Bt1Gc1330	Bt1G1509	785-1513	g3123142	200	246	6.50E-21	27	100	HYPOTHETICAL 26.1 KD PROTEIN IN PROX-MPRA INTERGENIC REGION [Escherichia coli]
992	Bt1Gc1330	Bt1G1510	3131-2570	g2632703	293	334	3.10E-30	39	99	(Z99106) yezH [Bacillus subtilis]
993	Bt1Gc1329	Bt1G1511	248-1	g2634266	198	210	4.20E-17	52	51	(Z99114) yoaS [Bacillus subtilis]
993	Bt1Gc1329	Bt1G1512	870-424	g2226251	389	398	5.10E-37	54	99	(Y14084) hypothetical protein [Bacillus subtilis]
993	Bt1Gc1329	Bt1G1513	1346-1	g3915359	112	182	1.40E-13	26	41	HYPOTHETICAL PROTEIN AF0433 [Archaeoglobus fulgidus]
993	Bt1Gc1329	Bt1G1514	2124-1438	g140681	663	672	4.70E-66	55	100	HYPOTHETICAL 25.2 KD PROTEIN IN LYSR-ARAE INTERGENIC REGION [Escherichia coli]
993	Bt1Gc1329	Bt1G1515	3488-2337	g462590	604	364	3.00E-54	36	100	NAD-DEPENDENT METHANOL DEHYDROGENASE (MEDH) [Bacillus sp.]
993	Bt1Gc1329	Bt1G1516	4532-3635	g1177018	880	867	1.00E-86	60	64	HYPOTHETICAL 48.3 KD PROTEIN IN KATB



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
994	Bt1Gc1331	Bt1G1517	35-481	g123762	442	462	8.40E-44	58	99	3'REGION [Bacillus subtilis] HUT OPERON POSITIVE REGULATORY PROTEIN [Bacillus subtilis]
994	Bt1Gc1331	Bt1G1518	588-2111	g123758	1825	1777	3.80E-183	70	100	HISTIDINE AMMONIA-LYASE (HISTIDASE) [Bacillus subtilis]
994	Bt1Gc1331	Bt1G1519	2129-2659	g1170427	721	646	2.70E-63	77	32	UROCANATE HYDRATASE (UROCANASE) (IMIDAZOLONEPROPIONATE HYDROLASE) [Bacillus subtilis]
995	Bt1Gc1332	Bt1G1520	1618-1	g1174516	1843	1846	1.80E-190	58	53	ISOLEUCYL-TRNA SYNTHETASE, MUPIROCIN RESISTANT (ISOLEUCINE--TRNA LIGASE) (ILERS) (MUPIROCIN RESISTANCE PROTEIN) [Staphylococcus aureus]
995	Bt1Gc1332	Bt1G1521	3494-97	g2388582	122	197	1.60E-14	27	27	(AC000098) Contains similarity to Rattus O-GlcNAc transferase (gb U76557). [Arabidopsis thaliana]
996	Bt1Gc1333	Bt1G1522	304-1008	g3608398	446	482	6.40E-46	46	100	(AF071085) putative glycosyl transferase [Enterococcus faecalis]
997	Bt1Gc1335	Bt1G1523	2008-200	g1865711	458	597	4.20E-58	35	100	(Y11477) endolysin [Bacteriophage Bastille]
998	Bt1Gc1334	Bt1G1524	1231-2361	g1730252	1640	1579	3.60E-162	86	100	GCPE PROTEIN HOMOLOG [Bacillus subtilis]
998	Bt1Gc1334	Bt1G1525	2842-2408	g1730119	496	472	7.30E-45	63	99	FERRIC UPTAKE REGULATION PROTEIN HOMOLOG 1 [Bacillus subtilis]
998	Bt1Gc1334	Bt1G1526	3686-2862	g3758894	356	373	2.30E-34	31	100	(Z71552) AdcB protein [Streptococcus pneumoniae]
998	Bt1Gc1334	Bt1G1527	4456-3720	g2415737	403	480	1.00E-45	39	100	(AB000617) YcdI [Bacillus subtilis]
998	Bt1Gc1334	Bt1G1528	4968-4643	g1731008	442	458	2.20E-43	81	37	HYPOTHETICAL 32.5 KD PROTEIN IN CCCA-SODA INTERGENIC REGION [Bacillus subtilis]
999	Bt1Gc1337	Bt1G1529	1283-150	g1926326	693	724	1.40E-71	39	100	(X98106) integrase [Bacteriophage phi1e]
1000	Bt1Gc1338	Bt1G1530	1307-456	g1864009	692	705	1.50E-69	49	100	(D85892) homologue of rdmC protein of Streptomyces purpuras [Escherichia coli]
1000	Bt1Gc1338	Bt1G1531	3071-1803	g3256593	453	504	3.00E-48	29	100	(AP000001) 428aa long hypothetical protein [Pyrococcus horikoshii]
1000	Bt1Gc1338	Bt1G1532	3233-1844	g2293169	113	162	5.20E-12	29	88	(AF008220) transcription regulator [Bacillus subtilis]
1000	Bt1Gc1338	Bt1G1533	3373-4292	g1176933	190	336	1.90E-30	26	100	HYPOTHETICAL 34.4 KD PROTEIN IN TRPA 3'REGION [Buchnera aphidicola]
1001	Bt1Gc1336	Bt1G1534	1707-907	g1706796	569	587	4.80E-57	42	100	FERRICHRONE TRANSPORT ATP-BINDING PROTEIN FHUC [Bacillus subtilis]
1001	Bt1Gc1336	Bt1G1535	2686-1694	g2126200	545	515	2.00E-49	38	100	hemin permease - Yersinia enterocolitica []

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1002	Bt1Gc1341	Bt1G1536	1768-1	g2612882	104	185	1.20E-13	27	48	(AF015825) NodB-like protein [Bacillus subtilis]
1002	Bt1Gc1341	Bt1G1537	962-1339	g116904	380	381	3.20E-35	57	99	COMA OPERON PROTEIN 2 [Bacillus subtilis]
1002	Bt1Gc1341	Bt1G1538	1768-1415	g2497400	183	231	2.50E-19	40	57	HYPOTHETICAL TRANSPOSASE-LIKE PROTEIN HI1721 [Haemophilus influenzae Rd]
1003	Bt1Gc1342	Bt1G1544	633-1	g2337818	742	765	6.50E-76	69	64	(Y13937) putative PlsX protein [Bacillus subtilis]
1003	Bt1Gc1342	Bt1G1545	1217-654	g2337817	685	688	9.40E-68	73	99	(Y13937) YlpC protein [Bacillus subtilis]
1003	Bt1Gc1342	Bt1G1546	2488-1315	g3914611	1516	1425	7.50E-146	76	57	ATP-DEPENDENT DNA HELICASE RECG [Bacillus subtilis]
1004	Bt1Gc1343	Bt1G1547	14-1712	g4585854	160	160	8.40E-12	37	48	(AJ237976) hypothetical protein [Streptomyces coelicolor]
1004	Bt1Gc1343	Bt1G1548	1295-363	g3183577	445	540	4.60E-52	36	100	HYPOTHETICAL OXIDOREDUCTASE IN NATB-RAP/INTERGENIC REGION (ORFC) [Bacillus subtilis]
1004	Bt1Gc1343	Bt1G1549	2820-3798	g4514327	756	706	1.20E-69	50	97	(AB013368) YkfB [Bacillus halodurans]
1005	Bt1Gc1344	Bt1G1550	531-335	g2879772	110	145	3.30E-10	37	45	(Y07640) putative mercury resistance operon regulatory protein (MerR) [Listeria monocytogenes]
1005	Bt1Gc1344	Bt1G1551	638-2439	g1403498	150	235	9.50E-20	33	77	(Z74020) hypothetical protein Rv1544 [Mycobacterium tuberculosis]
1005	Bt1Gc1344	Bt1G1552	2006-1474	g2983780	304	311	8.40E-28	41	99	(AE000736) hypothetical protein [Aquifex aeolicus]
1005	Bt1Gc1344	Bt1G1553	2575-2096	g1168677	270	264	8.00E-23	37	99	BKD OPERON TRANSCRIPTIONAL REGULATOR [Pseudomonas putida]
1005	Bt1Gc1344	Bt1G1554	2796-3085	g2650107	122	121	4.30E-07	35	33	(AE001068) conserved hypothetical protein [Archaeoglobus fulgidus]
1006	Bt1Gc1345	Bt1G1555	1-646	g421510	312	287	3.40E-25	34	50	kinB protein - Bacillus subtilis []
1006	Bt1Gc1345	Bt1G1556	3920-1017	g1169000	2254	2318	1.80E-240	47	88	MICROBIAL COLLAGENASE PRECURSOR (120 KD COLLAGENASE) [Clostridium perfringens]
1006	Bt1Gc1345	Bt1G1557	4805-4367	g421055	371	273	9.00E-24	49	99	hypothetical protein - Escherichia coli [Escherichia coli]
1007	Bt1Gc1346	Bt1G1558	662-87	g2635812	597	608	2.80E-59	61	99	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1007	Bt1Gc1346	Bt1G1559	2578-806	g466195	1491	1342	4.70E-137	53	100	SENSOR PROTEIN RESE [Bacillus subtilis]
1008	Bt1Gc1347	Bt1G1560	291-1136	g1763711	393	425	7.00E-40	32	100	(Z83337) similar to B. subtilis YcsE hypothetical protein [Bacillus subtilis]
1009	Bt1Gc1349	Bt1G1561	1-2439	g2635499	159	197	1.50E-14	29	22	(Z99119) similar to transcriptional regulator (AraC/XylS family) [Bacillus subtilis]
1010	Bt1Gc1350	Bt1G1562	1-444	g1168550	526	526	8.10E-50	65	16	MG(2+) TRANSPORT ATPASE, P-TYPE 2 [Salmonella typhimurium]
1010	Bt1Gc1350	Bt1G1563	520-1215	g2833487	517	486	2.40E-46	43	100	HYPOTHETICAL 25.5 KD PROTEIN SLR0014 [Synechocystis sp.]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1010	Bt1Gc1350	Bt1G1564	2488-1247	g1731071	1108	1080	2.70E-109	51	100	HYPOTHETICAL 47.0 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1011	Bt1Gc1351	Bt1G1565	1335-258	g3023206	606	667	1.60E-65	45	100	PUTATIVE 1-AMINOCYCLOPROPANE-1-CARBOXYLATE DEAMINASE (ACC DEAMINASE) [Escherichia coli]
1012	Bt1Gc1352	Bt1G1566	2586-3611	g4378847	370	482	6.40E-46	30	100	unknown [Zymomonas mobilis]
1012	Bt1Gc1352	Bt1G1567	4800-3647	g1651216	949	899	4.10E-90	47	61	(D88209) Pz-peptidase [Bacillus licheniformis]
1013	Bt1Gc1355	Bt1G1568	1258-1	g2632729	616	686	1.50E-67	35	74	(Z99106) ydaL [Bacillus subtilis]
1013	Bt1Gc1355	Bt1G1569	3711-2250	g2811053	70	113	8.30E-06	26	28	DI-/TRIPEPTIDE TRANSPORTER [Lactobacillus helveticus]
1014	Bt1Gc1353	Bt1G1570	139-1	g1731093	117	133	6.10E-09	55	25	HYPOTHETICAL 21.0 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1014	Bt1Gc1353	Bt1G1571	441-1355	g1731092	950	963	6.80E-97	60	100	HYPOTHETICAL OXIDOREDUCTASE IN ANSR-BMRU INTERGENIC REGION [Bacillus subtilis]
1014	Bt1Gc1353	Bt1G1572	2249-1419	g2635853	874	897	6.70E-90	60	100	(Z99121) alternate gene name: yvsB; similar to plant-metabolite dehydrogenase [Bacillus subtilis]
1014	Bt1Gc1353	Bt1G1573	3135-2312	g1731087	367	445	5.30E-42	38	72	HYPOTHETICAL 39.0 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1014	Bt1Gc1353	Bt1G1574	3135-2962	g1731091	111	68	0.046	36	77	HYPOTHETICAL 9.1 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1015	Bt1Gc1356	Bt1G1575	327-1302	g3123125	714	576	7.00E-56	45	100	HYPOTHETICAL 35.3 KD PROTEIN IN CSPC-NAP INTERGENIC REGION [Bacillus subtilis]
1015	Bt1Gc1356	Bt1G1576	1310-2007	g2829657	170	179	8.20E-14	25	100	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN DCP-NOHA INTERGENIC REGION [Escherichia coli]
1015	Bt1Gc1356	Bt1G1577	3298-2374	g1934656	793	653	4.80E-64	50	100	(U93876) cation transport protein YrdO [Bacillus subtilis]
1016	Bt1Gc1357	Bt1G1578	1840-993	g2632519	1138	1151	8.20E-117	71	88	(Z99105) similar to hypothetical proteins [Bacillus subtilis]
1017	Bt1Gc1358	Bt1G1579	1355-786	g2635871	258	270	1.90E-23	31	99	(Z99121) yvaF [Bacillus subtilis]
1017	Bt1Gc1358	Bt1G1580	2516-1638	g98344	465	386	9.50E-36	34	100	hypothetical protein (comG 5' region) - Bacillus subtilis (fragment) [Bacillus subtilis]
1018	Bt1Gc1348	Bt1G1581	1393-1	g1168885	788	784	6.30E-78	37	72	PUTATIVE CEL OPERON REGULATOR [Bacillus subtilis]
1018	Bt1Gc1348	Bt1G1582	2433-1479	g2689897	667	721	3.00E-71	45	89	(AF000792) outer surface protein, putative [Borrelia burgdorferi]
1019	Bt1Gc1359	Bt1G1583	682-1	g1731364	340	425	7.00E-40	39	55	HYPOTHETICAL PROTEASE IN ROCR-PURA INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1019	Bt1Gc1359	Bt1G1584	1575-895	g3236110	460	493	4.40E-47	43	100	(AJ001103) arcA [Lactococcus lactis]
1019	Bt1Gc1359	Bt1G1585	3097-1510	g4098082	212	321	7.30E-29	36	96	(U73336) anaerobic ribonucleotide reductase activator protein [Lactococcus lactis]
1019	Bt1Gc1359	Bt1G1586	3758-1504	g4098081	555	779	2.10E-77	30	80	(U73336) anaerobic ribonucleotide reductase [Lactococcus lactis]
1020	Bt1Gc1361	Bt1G1587	1-713	g3334448	528	552	2.40E-53	44	81	HYPOTHETICAL 33.2 KD PROTEIN IN GERAC-FHUC INTERGENIC REGION [Bacillus subtilis]
1020	Bt1Gc1361	Bt1G1588	2189-1560	g2619032	857	858	9.10E-86	74	100	(AF027868) YoaZ [Bacillus subtilis]
1020	Bt1Gc1361	Bt1G1589	3106-1887	g1881359	286	270	1.90E-23	32	100	(AB001488) SIMILAR TO YDFS GENE PRODUCT OF THIS ENTRY (YDFS_BACSU). [Bacillus subtilis]
1021	Bt1Gc1360	Bt1G1590	1-348	g135197	255	234	3.70E-19	51	28	TYROSYL-TRNA SYNTHETASE (TYROSINE--TRNA LIGASE) (TYRRS) []
1021	Bt1Gc1360	Bt1G1591	1838-663	g2612901	1700	1537	1.00E-157	85	100	(AF015825) NADH dehydrogenase-like protein [Bacillus subtilis]
1021	Bt1Gc1360	Bt1G1592	2249-1883	g2612900	490	217	1.40E-25	79	87	(AF015825) unknown [Bacillus subtilis]
1022	Bt1Gc1363	Bt1G1593	1309-95	g2632230	1196	1121	1.20E-113	55	100	(AJ222587) YkuI protein [Bacillus subtilis]
1022	Bt1Gc1363	Bt1G1594	2328-1567	g3915503	872	784	6.30E-78	66	100	HYPOTHETICAL OXIDOREDUCTASE IN CHEV-MOBA INTERGENIC REGION [Bacillus subtilis]
1022	Bt1Gc1363	Bt1G1595	2558-3415	g2632226	748	765	6.50E-76	50	100	(AJ222587) YkuE protein [Bacillus subtilis]
1023	Bt1Gc1367	Bt1G1596	1-289	g2632007	300	171	3.60E-12	63	22	(AJ002571) Ykba [Bacillus subtilis]
1023	Bt1Gc1367	Bt1G1597	2-1581	g1945096	952	778	2.70E-77	46	100	(D88802) S. lividans chloramphenicol resistance protein; P31141 (492) transmembrane [Bacillus subtilis]
1024	Bt1Gc1368	Bt1G1598	1-592	g3334472	653	681	5.20E-67	63	26	5-METHYLTETRAHYDROPTEROYLTRIGLUTAMATE--HOMOCYSTEINE METHYLTRANSFERASE (METHIONINE SYNTHASE, VITAMIN-B12 INDEPENDENT ISOZYME) (COBALAMIN-INDEPENDENT METHIONINE SYNTHASE) (SUPEROXIDE-INDUCIBLE PROTEIN 9) (SOI9) [Bacillus subtilis]
1024	Bt1Gc1368	Bt1G1599	1044-652	g112708	149	185	1.90E-14	30	99	COMPETENCE PROTEIN J (DNA-ENTRY NUCLEASE INHIBITOR) [Bacillus subtilis]
1024	Bt1Gc1368	Bt1G1600	1500-1063	g1171856	445	451	1.20E-42	57	99	DNA-ENTRY NUCLEASE (COMPETENCE-SPECIFIC NUCLEASE) []
1024	Bt1Gc1368	Bt1G1601	1-4096	g3549261	44	141	4.00E-08	24	10	(AF057019) interaptin [Dictyostelium discoideum]
1024	Bt1Gc1368	Bt1G1602	3783-2515	g1934780	699	685	2.00E-67	39	100	(Z93933) sporulation-specific ATP-dependent protein kinase

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1025	Bt1Gc1364	Bt1G1603	3364-1743	g3023642	173	256	9.40E-22	29	57	[Bacillus subtilis] DNA/PANTOTHENATE METABOLISM FLAVOPROTEIN HOMOLOG [Methanococcus jannaschii]
1025	Bt1Gc1364	Bt1G1604	3230-2417	g2634583	789	743	1.40E-73	56	100	(Z99115) similar to aminoglycoside N3'-acetyltransferase [Bacteriophage SPBc2]
1026	Bt1Gc1365	Bt1G1605	1-1233	g2619016	1716	1720	4.10E-177	78	83	(AF027868) aldehyde dehydrogenase [Bacillus subtilis]
1026	Bt1Gc1365	Bt1G1606	2017-1286	g1502421	545	564	1.30E-54	46	100	(U59433) 3-ketoacyl-acyl carrier protein reductase [Bacillus subtilis]
1027	Bt1Gc1369	Bt1G1607	1-1196	g1172784	1506	1457	3.10E-149	71	93	DIHYDROOROTASE (DHOASE) [Bacillus caldolyticus]
1027	Bt1Gc1369	Bt1G1608	1202-2291	g1705597	1364	1384	1.70E-141	71	100	CARBAMOYL-PHOSPHATE SYNTHASE, PYRIMIDINE-SPECIFIC, SMALL CHAIN (CARBAMOYL-PHOSPHATE SYNTHETASE GLUTAMINE CHAIN) [Bacillus caldolyticus]
1028	Bt1Gc1371	Bt1G1609	580-1	g4033437	344	362	3.30E-33	37	88	PYRUVATE KINASE (PK) [Thermococcus litoralis]
1028	Bt1Gc1371	Bt1G1610	411-2310	g4126634	243	382	2.50E-35	29	100	(AB016282) ORF48 [bacteriophage phi-105]
1028	Bt1Gc1371	Bt1G1611	4209-3436	g2633552	555	625	4.50E-61	48	100	(Z99110) alternate gene name: yidA; similar to 3-oxoacyl-acyl-carrier protein reductase [Bacillus subtilis]
1029	Bt1Gc1372	Bt1G1612	1-631	g1709569	584	604	7.50E-59	55	73	PANTOATE--BETA-ALANINE LIGASE (PANTOTHENATE SYNTHETASE) (PANTOATE ACTIVATING ENZYME) [Bacillus subtilis]
1029	Bt1Gc1372	Bt1G1613	631-1011	g1709570	479	479	1.30E-45	74	99	ASPARTATE 1-DECARBOXYLASE PRECURSOR (ASPARTATE ALPHA-DECARBOXYLASE) [Bacillus subtilis]
1029	Bt1Gc1372	Bt1G1614	1143-3912	g1706437	1653	1790	1.60E-184	39	99	PROBABLE ATP-DEPENDENT HELICASE DING HOMOLOG [Bacillus subtilis]
1030	Bt1Gc1374	Bt1G1615	1-496	g1001585	240	270	1.90E-23	35	50	(D64000) hypothetical protein [Synechocystis sp.]
1030	Bt1Gc1374	Bt1G1616	1-2437	g1806225	346	483	2.80E-45	27	58	(Z84498) hypothetical protein Rv1937 [Mycobacterium tuberculosis]
1030	Bt1Gc1374	Bt1G1617	2499-3027	g2633163	446	488	1.50E-46	51	99	(Z99108) yfiT [Bacillus subtilis]
1030	Bt1Gc1374	Bt1G1618	4000-3629	g2633601	384	396	8.30E-37	61	99	(Z99110) yjqa [Bacillus subtilis]
1030	Bt1Gc1374	Bt1G1619	3317-4900	g732119	112	241	2.20E-20	25	73	HYPOTHETICAL 39.8 KD PROTEIN IN OSMY-DEOC INTERGENIC REGION (O357) [Escherichia coli]
1031	Bt1Gc1373	Bt1G1620	1-834	g1817534	238	317	1.90E-28	30	60	(D78508) YfiJ [Bacillus subtilis]
1031	Bt1Gc1373	Bt1G1621	738-1462	g4753877	290	334	3.10E-30	35	100	(AL049754) putative two-component system regulator

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1031	Bt1Gc1373	Bt1G1622	3375-2427	g80141	1257	1255	7.80E-128	81	100	[Streptomyces coelicolor] beta-lactamase (EC 3.5.2.6) precursor - Bacillus sp. (strain 170) [Bacillus sp.]
1032	Bt1Gc1376	Bt1G1623	297-1172	g2828522	754	629	1.70E-61	51	100	HYPOTHETICAL 31.8 KD PROTEIN IN GABP-GUAA INTERGENIC REGION (ORFX) [Bacillus subtilis]
1032	Bt1Gc1376	Bt1G1624	1820-1174	g3222157	290	301	9.70E-27	31	100	hypothetical protein 2 - Bacillus megaterium [Bacillus megaterium]
1032	Bt1Gc1376	Bt1G1625	2651-1876	g3522984	471	527	1.10E-50	41	100	(D89936) sporulation-control protein (spo0M) [Bacillus subtilis]
1032	Bt1Gc1376	Bt1G1626	3450-2651	g2632489	374	385	1.20E-35	33	100	(Z99105) similar to protein kinase [Bacillus subtilis]
1033	Bt1Gc1375	Bt1G1628	1347-2228	g2181979	437	455	4.60E-43	37	100	(Z96072) hypothetical protein Rv2688c [Mycobacterium tuberculosis]
1033	Bt1Gc1375	Bt1G1629	3095-3814	g1592705	422	495	2.70E-47	45	86	(Y08559) Unknown [Bacillus subtilis]
1034	Bt1Gc1378	Bt1G1630	482-247	g2635194	74	118	3.30E-07	30	38	(Z99117) similar to hypothetical proteins [Bacillus subtilis]
1034	Bt1Gc1378	Bt1G1631	1047-247	g2983568	42	139	2.40E-08	23	21	(AE000722) hypothetical protein [Aquifex aeolicus]
1035	Bt1Gc1377	Bt1G1632	3582-1756	g1934835	1444	1460	1.50E-149	45	100	(Z93940) asparagine synthetase [Bacillus subtilis]
1036	Bt1Gc1379	Bt1G1633	876-1	g1648861	1143	1114	6.80E-113	76	67	(Z81356) UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Bacillus subtilis]
1036	Bt1Gc1379	Bt1G1634	1774-1678	g2636203	101	65	0.094	67	42	(Z99122) ywzB [Bacillus subtilis]
1037	Bt1Gc1383	Bt1G1635	724-1	g2495457	842	713	2.10E-70	66	61	HYPOTHETICAL 41.5 KD PROTEIN IN AMHX-AMYE INTERGENIC REGION [Bacillus subtilis]
1037	Bt1Gc1383	Bt1G1636	1152-1622	g1781103	310	320	9.40E-29	41	99	(Z83864) menG [Mycobacterium tuberculosis]
1038	Bt1Gc1382	Bt1G1637	1766-1	g730275	1360	1394	1.50E-142	50	64	PENICILLIN-BINDING PROTEINS 1A/1B (PBP1) [Bacillus subtilis]
1038	Bt1Gc1382	Bt1G1638	2432-1815	g732244	645	675	2.30E-66	63	100	HYPOTHETICAL 24.0 KD PROTEIN IN PONA-COTD INTERGENIC REGION (ORF1) [Bacillus subtilis]
1038	Bt1Gc1382	Bt1G1639	2590-3543	g1730944	688	738	4.80E-73	46	100	HYPOTHETICAL 38.5 KD PROTEIN IN PONA-COTD INTERGENIC REGION (ORFX) [Bacillus subtilis]
1039	Bt1Gc1380	Bt1G1640	933-1	g790865	93	151	7.60E-11	28	74	(L36381) orfA; putative [Neisseria gonorrhoeae]
1040	Bt1Gc1370	Bt1G1641	299-2116	g1723607	1389	1303	6.40E-133	46	100	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN 2 IN GLVBC 3'REGION [Bacillus subtilis]
1041	Bt1Gc1384	Bt1G1642	546-1931	g2634220	1758	1500	8.50E-154	73	100	(Z99113) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1041	Bt1Gc1384	Bt1G1643	2698-927	g1673400	199	314	4.00E-28	29	82	(Z82044) hypothetical 28.3 kd protein [Bacillus subtilis]
1042	Bt1Gc1381	Bt1G1644	1-259	g133395	389	405	9.20E-38	95	27	DNA-DIRECTED RNA POLYMERASE ALPHA CHAIN

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1042	Bt1Gc1381	Bt1G1645	298-657	g132718	528	528	8.50E-51	88	99	(TRANSCRIPTASE ALPHA CHAIN) (RNA POLYMERASE ALPHA SUBUNIT) [Bacillus subtilis]
1042	Bt1Gc1381	Bt1G1646	761-1603	g3915966	831	868	8.00E-87	60	100	50S RIBOSOMAL PROTEIN L17 (BL15) [Bacillus subtilis]
1042	Bt1Gc1381	Bt1G1647	1621-2405	g1644203	852	803	6.20E-80	61	95	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YBXA [Bacillus subtilis]
1043	Bt1Gc1388	Bt1G1648	1-1189	g3763917	222	327	1.70E-29	37	76	(D64126) unknown [Bacillus subtilis]
1043	Bt1Gc1388	Bt1G1649	1305-1811	g114199	187	186	1.50E-14	29	99	hypothetical protein [Arabidopsis thaliana]
1043	Bt1Gc1388	Bt1G1650	1844-2955	g2634900	150	138	1.80E-09	42	99	SHIKIMATE KINASE (SK) [Erwinia chrysanthemi]
1043	Bt1Gc1388	Bt1G1651	3872-3575	g1731039	247	270	1.90E-23	47	38	(Z99116) yqzE [Bacillus subtilis]
1044	Bt1Gc1386	Bt1G1652	993-313	g3599371	432	411	2.10E-38	41	100	HYPOTHETICAL 30.8 KD PROTEIN IN SINI-GCVT INTERGENIC REGION [Bacillus subtilis]
1045	Bt1Gc1385	Bt1G1653	1-4206	g2982194	1392	1485	6.80E-150	38	16	(AF082668) CsrR [Streptococcus pyogenes]
1045	Bt1Gc1385	Bt1G1654	3617-4206	g2497382	912	882	2.60E-88	91	45	(AF007865) bacitracin synthetase 1; BacA [Bacillus licheniformis]
1046	Bt1Gc1390	Bt1G1655	225-1	g2226185	210	210	5.60E-17	52	22	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1046	Bt1Gc1390	Bt1G1656	941-426	g1176283	89	136	2.90E-09	27	99	(Y14081) hypothetical protein [Bacillus subtilis]
1047	Bt1Gc1387	Bt1G1657	421-1	g416876	248	272	1.10E-23	33	48	HYPOTHETICAL 18.8 KD PROTEIN IN GNTR-GGT INTERGENIC REGION (O162) [Escherichia coli]
1047	Bt1Gc1387	Bt1G1658	3305-2440	g3915984	157	238	5.60E-20	25	73	DIHYDRODIPICOLINATE SYNTHASE (DHDPS) (VEGETATIVE PROTEIN 81) (VEG81) [Bacillus subtilis]
1048	Bt1Gc1391	Bt1G1659	889-1	g1750108	634	628	2.20E-61	48	86	HYPOTHETICAL 39.5 KD OXIDOREDUCTASE IN FIMT 3 REGION (DADA*) (ORFZ) [Pseudomonas aeruginosa]
1048	Bt1Gc1391	Bt1G1660	1656-2630	g2226235	1119	1045	1.40E-105	64	100	(U66480) YnbA [Bacillus subtilis]
1048	Bt1Gc1391	Bt1G1661	3253-2695	g1881327	686	304	4.60E-27	73	43	(Y14083) Hypothetical protein [Bacillus subtilis]
1049	Bt1Gc1392	Bt1G1662	1-1531	g730002	595	227	1.90E-25	35	62	(AB001488) SIMILAR TO YDDS_BACSU. [Bacillus subtilis]
1049	Bt1Gc1392	Bt1G1663	1977-2709	g4894306	428	508	1.10E-48	46	37	METHYL-ACCEPTING CHEMOTAXIS PROTEIN MCPA (H1) [Bacillus subtilis]
1050	Bt1Gc1393	Bt1G1664	1760-968	g1711477	1303	1319	1.30E-134	99	100	(AF065404) pXO1-90 [Bacillus anthracis]
1050	Bt1Gc1393	Bt1G1665	3337-2057	g1553038	1365	1388	6.30E-142	64	100	STAGE 0 SPORULATION PROTEIN A [Bacillus thuringiensis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1050	Bt1Gc1393	Bt1G1666	3793-3523	g4106732	335	364	2.00E-33	80	73	(AF055996) RecN homolog [Bacillus anthracis]
1051	Bt1Gc1395	Bt1G1667	434-1	g773349	415	375	1.40E-34	56	44	(U20445) BirA protein [Bacillus subtilis]
1051	Bt1Gc1395	Bt1G1668	1615-422	g1709578	819	878	6.90E-88	44	100	POLY(A) POLYMERASE (PAP) [Bacillus subtilis]
1051	Bt1Gc1395	Bt1G1669	2597-1617	g1730929	1030	978	1.80E-98	60	86	HYPOTHETICAL 42.0 KD PROTEIN IN DAPB-PAPS INTERGENIC REGION [Bacillus subtilis]
1052	Bt1Gc1396	Bt1G1670	573-1496	g2293314	1032	914	1.10E-91	67	100	(AF008220) putative cysteine synthase [Bacillus subtilis]
1053	Bt1Gc1398	Bt1G1671	1-1529	g585920	2283	2292	1.00E-237	87	43	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Bacillus subtilis]
1054	Bt1Gc1397	Bt1G1672	331-1362	g2507429	1411	1286	4.00E-131	76	100	PHENYLALANYL-TRNA SYNTHETASE ALPHA CHAIN (PHENYLALANINE--TRNA LIGASE ALPHA CHAIN) (PHERS) [Bacillus subtilis]
1054	Bt1Gc1397	Bt1G1673	1384-3108	g3123292	1865	1780	1.80E-183	65	71	PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PHERS) [Bacillus subtilis]
1055	Bt1Gc1394	Bt1G1674	Jan-58	g2634035	2375	2321	1.80E-242	73	91	(Z99112) initiation factor IF-2 [Bacillus subtilis]
1055	Bt1Gc1394	Bt1G1675	1958-2233	g418463	306	306	2.90E-27	61	99	HYPOTHETICAL 10.7 KD PROTEIN IN INF8-RPSO INTERGENIC REGION (ORF5) [Bacillus subtilis]
1055	Bt1Gc1394	Bt1G1676	2252-2602	g418464	482	482	6.40E-46	80	99	RIBOSOME-BINDING FACTOR A (P15B PROTEIN) [Bacillus subtilis]
1055	Bt1Gc1394	Bt1G1677	2692-3609	g3183559	919	955	4.80E-96	59	100	TRNA PSEUDOURIDINE SYNTHASE B (TRNA PSEUDOURIDINE 55 SYNTHASE) (PSI55 SYNTHASE) (PSEUDOURIDYLATE SYNTHASE) (URACIL HYDROLYASE) [Bacillus subtilis]
1056	Bt1Gc1399	Bt1G1678	1-1847	g2493270	397	409	3.50E-38	41	60	HYPOTHETICAL 41.2 KD PROTEIN CY277.09 [Mycobacterium tuberculosis]
1056	Bt1Gc1399	Bt1G1679	825-1934	g3724051	660	692	3.60E-68	37	100	(AJ000758) hypothetical protein [Bacillus megaterium]
1056	Bt1Gc1399	Bt1G1680	2015-2611	g2226254	150	189	7.10E-15	30	99	(Y14084) hypothetical protein [Bacillus subtilis]
1056	Bt1Gc1399	Bt1G1681	3104-3391	g141235	193	208	6.90E-17	51	99	HYPOTHETICAL PROTEIN IN PDHA 5'REGION (ORF1) [Bacillus stearothermophilus]
1057	Bt1Gc1401	Bt1G1682	1-314	g225559	502	518	9.80E-50	98	22	ORF IS231C [Bacillus thuringiensis]
1057	Bt1Gc1401	Bt1G1683	1621-536	g399057	785	821	7.60E-82	45	100	3-DEHYDROQUINATE SYNTHASE [Bacillus subtilis]
1057	Bt1Gc1401	Bt1G1684	2134-1686	g399058	604	432	2.90E-42	83	40	CHORISMATE SYNTHASE (5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE PHOSPHOLYASE) (VEGETATIVE PROTEIN 216)



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1058	Bt1Gc1400	Bt1G1685	316-1	g2633967	392	364	2.00E-33	77	32	(VEG216) [Bacillus subtilis] (Z99112) signal recognition particle (docking protein) [Bacillus subtilis]
1058	Bt1Gc1400	Bt1G1686	3924-341	g2633966	2758	2653	5.60E-276	53	90	(Z99112) chromosome segregation SMC protein homolog [Bacillus subtilis]
1059	Bt1Gc1403	Bt1G1687	186-1890	g129179	233	426	5.50E-40	24	100	OLIGOPEPTIDE-BINDING PROTEIN OPPA PRECURSOR [Bacillus subtilis]
1059	Bt1Gc1403	Bt1G1688	174-1852	g3257217	467	669	9.70E-66	33	87	(AP000003) 597aa long hypothetical oligopeptide binding protein APPA [Pyrococcus horikoshii]
1059	Bt1Gc1403	Bt1G1689	2005-2925	g2983140	332	366	1.20E-33	40	65	(AE000692) transporter (OppBC family) [Aquifex aeolicus]
1059	Bt1Gc1403	Bt1G1690	2693-2925	g225559	364	380	4.10E-35	97	16	ORF IS231C [Bacillus thuringiensis]
1060	Bt1Gc1402	Bt1G1691	50-1816	g1945644	1775	1595	7.30E-164	59	100	(Z94043) hypothetical protein [Bacillus subtilis]
1060	Bt1Gc1402	Bt1G1692	1893-2573	g586808	433	354	5.10E-35	40	100	HYPOTHETICAL 27.2 KD SENSORY TRANSDUCTION PROTEIN IN ROCP-PURA INTERGENIC REGION [Bacillus subtilis]
1060	Bt1Gc1402	Bt1G1693	2844-3643	g130130	408	418	3.90E-39	38	47	ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN PHOR [Bacillus subtilis]
1061	Bt1Gc1405	Bt1G1694	551-252	g420808	408	351	4.90E-32	74	99	hypothetical 12K protein - Thermus aquaticus [Thermus aquaticus]
1061	Bt1Gc1405	Bt1G1695	762-2556	g2127359	171	206	1.10E-16	38	37	hypothetical protein 3 - Clostridium perfringens [Clostridium perfringens]
1061	Bt1Gc1405	Bt1G1696	3318-2092	g1075793	1316	1285	5.20E-131	67	100	aspartate kinase (EC 2.7.2.4) II precursor - Bacillus sp. (strain MGA3) [Bacillus sp.]
1062	Bt1Gc1404	Bt1G1697	930-1967	g1881258	928	653	4.80E-64	52	100	(AB001488) FUNCTION UNKNOWN, SIMILAR PRODUCTS IN SYNECHOCYSTIS AND H. INFLUENZAE. [Bacillus subtilis]
1062	Bt1Gc1404	Bt1G1698	2977-2021	g126054	1132	1104	7.80E-112	66	100	L-LACTATE DEHYDROGENASE X []
1062	Bt1Gc1404	Bt1G1699	3274-3917	g2619006	731	745	8.60E-74	67	48	(AF027868) putative L-amino acid oxidase precursor [Bacillus subtilis]
1063	Bt1Gc1407	Bt1G1700	47-808	g1731129	785	317	3.00E-31	56	100	HYPOTHETICAL 27.6 KD PROTEIN IN BLTR-SPOIIC INTERGENIC REGION [Bacillus subtilis]
1063	Bt1Gc1407	Bt1G1701	2645-1635	g400209	337	427	4.30E-40	40	100	MEMBRANE-BOUND PROTEIN LYTR [Bacillus subtilis]
1063	Bt1Gc1407	Bt1G1702	2993-3410	g2226252	293	321	7.30E-29	47	42	(Y14084) hypothetical protein [Bacillus subtilis]
1064	Bt1Gc1406	Bt1G1703	503-18	g971590	576	578	4.30E-56	60	99	(Z50141) dihydrofolate reductase [Listeria monocytogenes]
1064	Bt1Gc1406	Bt1G1704	1473-526	g136612	1217	1220	4.00E-124	70	100	THYMIDYLATE SYNTHASE (TS) [Artificial gene]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1064	Bt1Gc1406	Bt1G1705	507-3114	g3261758	311	456	3.60E-43	24	92	(Z95210) hypothetical protein Rv0907 [Mycobacterium tuberculosis]
1064	Bt1Gc1406	Bt1G1706	5156-4497	g1652645	365	390	3.60E-36	34	100	(D90907) phosphoglycolate phosphatase [Synechocystis sp.]
1064	Bt1Gc1406	Bt1G1707	5367-5260	g1730930	122	126	3.40E-08	53	17	HYPOTHETICAL 23.7 KD PROTEIN IN ILVD-THYB INTERGENIC REGION [Bacillus subtilis]
1065	Bt1Gc1408	Bt1G1708	224-1	g2072371	332	346	1.60E-31	92	24	(Y09212) putative malate oxidoreductase [Bacillus cereus]
1065	Bt1Gc1408	Bt1G1709	1685-249	g2072370	2378	2200	5.60E-228	99	100	(Y09212) Aspartate ammonia-lyase [Bacillus cereus]
1065	Bt1Gc1408	Bt1G1710	2777-1713	g2072369	1733	1512	4.50E-155	99	100	(Y09212) proton /sodium-glutamate symport protein [Bacillus cereus]
1066	Bt1Gc1409	Bt1G1711	1354-557	g3688282	793	722	2.40E-71	56	100	(AJ010739) pyrroline-5-carboxylate reductase [Clostridium sticklandii]
1066	Bt1Gc1409	Bt1G1712	2046-3152	g3183548	1112	1010	7.10E-102	60	100	GLUTAMATE 5-KINASE (GAMMA-GLUTAMYL KINASE) (GK) [Bacillus subtilis]
1066	Bt1Gc1409	Bt1G1713	3166-4410	g2632033	1401	1401	2.60E-143	65	100	(AJ002571) gamma-glutamylphosphate reductase [Bacillus subtilis]
1067	Bt1Gc1410	Bt1G1714	496-56	g2621778	176	203	2.30E-16	32	99	(AE000848) conserved protein [Methanobacterium thermoautotrophicum]
1067	Bt1Gc1410	Bt1G1715	761-1976	g2443252	626	586	6.10E-57	33	100	(D86417) YfmI [Bacillus subtilis]
1067	Bt1Gc1410	Bt1G1716	3498-2035	g1813489	1092	1156	2.40E-117	50	100	(U64312) amidase [Bacillus firmus]
1068	Bt1Gc1411	Bt1G1717	2394-3543	g1881268	908	917	5.10E-92	49	75	(AB001488) ATP-DEPENDENT RNA HELICASE DEAD HOMOLOG. [Bacillus subtilis]
1069	Bt1Gc1413	Bt1G1718	981-238	g4914622	800	807	2.30E-80	59	100	(AJ009627) pyruvate-formate lyase activating enzyme [Listeria monocytogenes]
1069	Bt1Gc1413	Bt1G1719	1979-1048	g129879	1093	1123	7.60E-114	68	41	FORMATE ACETYLTRANSFERASE 1 (PYRUVATE FORMATE-LYASE 1) [Escherichia coli]
1070	Bt1Gc1412	Bt1G1720	1398-904	g2618995	493	503	3.80E-48	53	99	(AF027868) putative alanine acetyl transferase [Bacillus subtilis]
1070	Bt1Gc1412	Bt1G1721	2875-1601	g3025117	1603	1225	1.20E-124	73	100	HYPOTHETICAL 45.7 KD PROTEIN IN MUTT-GSIB INTERGENIC REGION [Bacillus subtilis]
1071	Bt1Gc1415	Bt1G1722	1352-2647	g4584200	2122	2025	2.00E-209	98	90	(AJ000394) transcriptional regulatory protein CelR [Bacillus cereus]
1072	Bt1Gc1416	Bt1G1723	1-1293	g2145380	845	891	2.90E-89	40	74	(Y09476) YisO [Bacillus subtilis]
1072	Bt1Gc1416	Bt1G1724	3570-3169	g3005554	105	162	1.60E-11	24	41	(AF047044) putative transposase [Anabaena PCC7120]
1073	Bt1Gc1414	Bt1G1725	1146-1	g2633729	1389	1392	2.40E-142	68	96	(Z99111) similar to aspartate aminotransferase [Bacillus

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1073	Bt1Gc1414	Bt1G1726	1242-2021	g2633728	758	778	2.70E-77	55	100	subtilis] (Z99111) similar to hypothetical proteins [Bacillus subtilis]
1073	Bt1Gc1414	Bt1G1727	2307-3497	g2633727	1027	1027	1.10E-103	54	100	(Z99111) ykrT [Bacillus subtilis]
1073	Bt1Gc1414	Bt1G1728	3509-4389	g3183453	625	670	7.60E-66	58	64	HYPOTHETICAL 38.9 KD PROTEIN IN DAT-SPO0E INTERGENIC REGION [Bacillus subtilis]
1073	Bt1Gc1414	Bt1G1729	4389-4288	g225559	143	164	2.40E-11	97	7	ORF IS231C [Bacillus thuringiensis]
1074	Bt1Gc1417	Bt1G1730	1987-926	g3915578	772	791	1.10E-78	42	100	HYPOTHETICAL OXIDOREDUCTASE IN FHU-D-OPUD INTERGENIC REGION [Bacillus subtilis]
1074	Bt1Gc1417	Bt1G1731	4314-2231	g2558947	2407	2302	8.80E-239	67	86	(AF024713) ParC [Bacillus subtilis]
1075	Bt1Gc1418	Bt1G1732	473-144	g2634056	254	260	2.10E-22	46	99	(Z99112) ymf [Bacillus subtilis]
1075	Bt1Gc1418	Bt1G1733	1738-782	g1934814	1058	918	4.00E-92	63	100	(Z93937) unknown [Bacillus subtilis]
1075	Bt1Gc1418	Bt1G1734	2782-1739	g1934813	1069	949	2.10E-95	59	100	(Z93937) unknown [Bacillus subtilis]
1075	Bt1Gc1418	Bt1G1735	4276-2793	g1934812	1711	1625	4.80E-167	68	97	(Z93937) unknown [Bacillus subtilis]
1076	Bt1Gc1419	Bt1G1736	1102-2	g1730981	1340	1350	6.70E-138	68	100	HYPOTHETICAL 41.0 KD PROTEIN IN NUCB-AROD INTERGENIC REGION [Bacillus subtilis]
1076	Bt1Gc1419	Bt1G1737	1627-1112	g1730980	638	569	3.80E-55	67	99	HYPOTHETICAL 20.1 KD PROTEIN IN NUCB-AROD INTERGENIC REGION [Bacillus subtilis]
1076	Bt1Gc1419	Bt1G1738	3009-2221	g729365	577	585	7.80E-57	43	100	PHOSPHATIDYL-SERINE DECARBOXYLASE PROENZYME [Bacillus subtilis]
1076	Bt1Gc1419	Bt1G1739	4000-3290	g133282	1169	1169	1.00E-118	100	100	RNA POLYMERASE SIGMA-28 FACTOR PRECURSOR [Bacillus thuringiensis]
1077	Bt1Gc1420	Bt1G1740	1-1792	g730820	157	213	8.80E-17	42	22	SPORE COAT POLYSACCHARIDE BIOSYNTHESIS PROTEIN SPSK [Bacillus subtilis]
1077	Bt1Gc1420	Bt1G1741	412-1216	g2633526	982	1004	3.10E-101	70	100	(Z99110) similar to enoyl-acyl-carrier protein reductase [Bacillus subtilis]
1077	Bt1Gc1420	Bt1G1742	1497-1897	g2633527	94	101	4.10E-05	27	60	(Z99110) yjbX [Bacillus subtilis]
1077	Bt1Gc1420	Bt1G1743	2542-2895	g2633533	274	175	2.20E-13	47	99	(Z99110) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1078	Bt1Gc1421	Bt1G1744	953-75	g1177035	796	714	1.70E-70	52	100	PROBABLE ABC TRANSPORTER PERMEASE PROTEIN IN SODA-COMGA INTERGENIC REGION (ORF73) [Bacillus subtilis]
1078	Bt1Gc1421	Bt1G1745	2040-959	g1177034	699	539	5.80E-52	49	100	PROBABLE ABC TRANSPORTER PERMEASE PROTEIN IN SODA-COMGA INTERGENIC REGION (ORF72) [Bacillus subtilis]
1078	Bt1Gc1421	Bt1G1746	2811-1894	g1177033	586	581	2.10E-56	43	100	PROBABLE ABC TRANSPORTER BINDING PROTEIN

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
IN SODA-COMGA INTERGENIC REGION PRECURSOR (ORF108) [Bacillus subtilis]										
1078	Bt1Gc1421	Bt1G1747	3342-3566	g2622470	85	106	4.50E-06	29	99	(AE000899) unknown [Methanobacterium thermoautotrophicum]
1078	Bt1Gc1421	Bt1G1748	3584-3895	g2498377	332	332	3.30E-29	60	15	FERROUS IRON TRANSPORT PROTEIN B HOMOLOG [Methanococcus jannaschii]
1079	Bt1Gc1422	Bt1G1749	1-363	g1575061	173	218	6.00E-18	34	53	(U57060) ScdA [Staphylococcus aureus]
1079	Bt1Gc1422	Bt1G1750	581-2992	g1171661	2412	2397	7.50E-249	57	100	NITRITE REDUCTASE (NAD(P)H) [Bacillus subtilis]
1079	Bt1Gc1422	Bt1G1751	3002-3319	g1171662	289	303	5.90E-27	49	99	ASSIMILATORY NITRITE REDUCTASE (NAD(P)H) SMALL SUBUNIT [Bacillus subtilis]
1079	Bt1Gc1422	Bt1G1752	3384-3709	g2462960	225	266	4.90E-23	46	42	(AJ000974) putative S-adenosyl L-methionine: uroporphyrinogen III methyltransferase [Bacillus subtilis]
1080	Bt1Gc1423	Bt1G1753	469-1458	g135188	1325	1325	3.00E-135	74	100	TRYPTOPHANYL-TRNA SYNTHETASE (TRYPTOPHAN--TRNA LIGASE) (TRPRS) [Bacillus subtilis]
1080	Bt1Gc1423	Bt1G1754	2488-1745	g2633495	975	985	3.20E-99	70	100	(Z99110) yjbA [Bacillus subtilis]
1080	Bt1Gc1423	Bt1G1755	3420-2647	g2633489	355	379	5.20E-35	36	100	(Z99110) yjaZ [Bacillus subtilis]
1081	Bt1Gc1424	Bt1G1756	1731-1	g2621399	169	135	1.20E-12	23	94	(AE000819) succinoglycan biosynthesis transport protein [Methanobacterium thermoautotrophicum]
1081	Bt1Gc1424	Bt1G1757	1728-1	g2650029	95	176	1.20E-12	27	42	(AE001064) polysaccharide biosynthesis protein, putative [Archaeoglobus fulgidus]
1081	Bt1Gc1424	Bt1G1758	2438-1909	g2454556	346	344	2.70E-31	44	99	(AF015609) unknown [Bacillus subtilis]
1081	Bt1Gc1424	Bt1G1759	3818-2410	g1495280	91	166	2.00E-12	22	100	(Z71928) hypothetical protein [Bacillus subtilis]
1081	Bt1Gc1424	Bt1G1760	3789-3049	g1894741	297	313	5.20E-28	31	100	(Z92952) product similar to Staphylococcus aureus CapA protein [Bacillus subtilis]
1082	Bt1Gc1425	Bt1G1761	2661-1516	g1730908	1064	1077	5.70E-109	54	100	HYPOTHETICAL 43.6 KD PROTEIN IN CPSD-METB INTERGENIC REGION [Bacillus subtilis]
1083	Bt1Gc1427	Bt1G1762	545-3017	g2492945	632	703	2.40E-69	32	96	ARGININOSUCCINATE LYASE (ARGINOSUCCINASE) (ASAL) [Methanococcus jannaschii]
1084	Bt1Gc1428	Bt1G1763	1-990	g2635928	1167	958	2.30E-96	62	79	(Z99121) similar to maltodextrin transport system permease [Bacillus subtilis]
1084	Bt1Gc1428	Bt1G1764	994-1839	g1945713	860	823	4.70E-82	58	100	(Z94043) hypothetical protein [Bacillus subtilis]
1084	Bt1Gc1428	Bt1G1765	2112-3110	g2656094	554	623	7.30E-61	40	100	(L21856) repressor protein [Streptococcus pneumoniae]
1084	Bt1Gc1428	Bt1G1767	4311-3130	g98327	1240	1280	1.70E-130	61	100	hisC homolog - Bacillus subtilis []
1085	Bt1Gc1431	Bt1G1768	411-1482	g728897	1374	1292	9.30E-132	75	100	PHOSPHO-2-DEHYDRO-3-DEOXYHEPTONATE

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1085	Bt1Gc1431	Bt1G1769	1770-2877	g399058	1430	683	1.60E-115	75	100	ALDOLASE (PHOSPHO-2-KETO-3-DEOXYHEPTONATE ALDOLASE) (DAHP SYNTHETASE) (3-DEOXY-D-ARABINO-HEPTULOSONATE 7-PHOSPHATE SYNTHASE) / CHORISMATE MUTASE [Bacillus subtilis]
1085	Bt1Gc1431	Bt1G1770	2961-4019	g3123224	869	903	1.60E-90	49	100	CHORISMATE SYNTHASE (5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE PHOSPHOLYASE) (VEGETATIVE PROTEIN 216) (VEG216) [Bacillus subtilis]
1086	Bt1Gc1435	Bt1G1771	115-321	g2635175	149	171	5.80E-13	49	99	HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (IMIDAZOLE ACETOL-PHOSPHATE TRANSAMINASE) [Bacillus subtilis]
1086	Bt1Gc1435	Bt1G1772	1031-368	g2635176	274	198	7.90E-16	31	100	yrzA [Bacillus subtilis]
1086	Bt1Gc1435	Bt1G1773	2895-1148	g2635177	1224	1293	7.30E-132	43	100	yrzS [Bacillus subtilis]
1086	Bt1Gc1435	Bt1G1774	4295-3822	g3183527	615	624	5.70E-61	80	99	similar to penicillin-binding protein [Bacillus subtilis]
1086	Bt1Gc1435	Bt1G1775	5162-4578	g2635179	688	706	1.20E-69	69	92	TRANSCRIPTION ELONGATION FACTOR GRE A (TRANSCRIPT CLEAVAGE FACTOR GRE A) (GENERAL STRESS PROTEIN 20M) (GSP20M) [Bacillus subtilis]
1087	Bt1Gc1434	Bt1G1776	588-16	g1108679	324	309	1.40E-27	38	99	uridine kinase [Bacillus subtilis]
1087	Bt1Gc1434	Bt1G1777	3434-2154	g114169	1429	1461	1.20E-149	66	100	TRANSCRIPTIONAL REPRESSOR BM3R1 [Bacillus megaterium]
1087	Bt1Gc1434	Bt1G1778	4402-3464	g136591	762	779	2.10E-77	48	84	PUTATIVE 3-PHOSPHOSHAKIMATE 1-CARBOXYVINYLTRANSFERASE (5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE) [Bacillus subtilis]
1088	Bt1Gc1429	Bt1G1779	818-1	g2339999	509	529	6.70E-51	38	77	PREPHENATE DEHYDROGENASE (PDH) []
1088	Bt1Gc1429	Bt1G1780	1002-1361	g2339997	348	349	7.90E-32	51	99	YlbC protein [Bacillus subtilis]
1088	Bt1Gc1429	Bt1G1781	2415-1417	g2314402	1398	1379	5.60E-141	75	100	YlbA protein [Bacillus subtilis]
1088	Bt1Gc1429	Bt1G1782	3548-2640	g2339996	752	667	1.60E-65	47	100	(AE000629) aliphatic amidase (aimE) [Helicobacter pylori 26695]
1088	Bt1Gc1429	Bt1G1783	3964-3632	g1841880	331	360	5.40E-33	62	99	CtaG protein [Bacillus subtilis]
										(D70843) cytochrome c oxidase subunit IV [Bacillus stearothermophilus]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1088	Bt1Gc1429	Bt1G1784	4334-3968	g2339994	544	470	1.20E-44	81	59	(Z98682) cytochrome c oxidase subunit III, aa3 type [Bacillus subtilis]
1089	Bt1Gc1433	Bt1G1785	1-213	g399378	206	237	5.80E-20	61	35	SOS REGULATORY PROTEIN LEXA/DINR [Bacillus subtilis]
1089	Bt1Gc1433	Bt1G1786	1974-1660	g730644	155	173	3.50E-13	32	99	30S RIBOSOMAL PROTEIN S14 HOMOLOG [Bacillus sphaericus]
1089	Bt1Gc1433	Bt1G1787	3924-2593	g121357	2288	2288	2.70E-237	98	100	GLUTAMINE SYNTHETASE (GLUTAMATE--AMMONIA LI GASE) [Bacillus cereus]
1090	Bt1Gc1426	Bt1G1788	1165-80	g2689897	636	684	2.50E-67	41	100	(AE000792) outer surface protein, putative [Borrelia burgdorferi]
1090	Bt1Gc1426	Bt1G1789	2722-1415	g732331	767	628	2.20E-61	41	100	HYPOTHETICAL 47.6 KD PROTEIN IN EPR-GALK INTERGENIC REGION [Bacillus subtilis]
1090	Bt1Gc1426	Bt1G1790	3111-2806	g1172709	247	257	4.40E-22	50	99	PTS SYSTEM, CELLOBIOSE-SPECIFIC IIB COMPONENT (IIB-CEL) (CELLOBIOSE-PERMEASE IIB COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, B COMPONENT) [Bacillus subtilis]
1090	Bt1Gc1426	Bt1G1791	3439-3110	g1172708	260	280	1.60E-24	50	99	PTS SYSTEM, CELLOBIOSE-SPECIFIC IIA COMPONENT (EIIA-CEL) (CELLOBIOSE-PERMEASE IIA COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, A COMPONENT) (EII-CEL) [Bacillus subtilis]
1090	Bt1Gc1426	Bt1G1792	4057-3538	g1168885	257	276	3.40E-23	33	27	PUTATIVE CEL OPERON REGULATOR [Bacillus subtilis]
1091	Bt1Gc1430	Bt1G1793	461-3	g2497629	539	377	8.50E-35	66	99	LIPOPROTEIN SIGNAL PEPTIDASE (PROLIPOPROTEIN SIGNAL PEPTIDASE) (SIGNAL PEPTIDASE-II) (SPASE II) [Bacillus subtilis]
1091	Bt1Gc1430	Bt1G1794	984-539	g3183580	97	97	4.00E-05	29	99	HYPOTHETICAL 16.8 KD PROTEIN IN ILES-LSP INTERGENIC REGION (ORF-T) [Bacillus subtilis]
1091	Bt1Gc1430	Bt1G1795	3836-1074	g3123287	3685	3712	0	74	100	ISOLEUCYL-TRNA SYNTHETASE (ISOLEUCINE--TRNA LI GASE) (ILERS) [Bacillus subtilis]
1091	Bt1Gc1430	Bt1G1796	4689-4192	g1518680	594	615	5.10E-60	70	99	(U60901) minicell-associated protein DivIVA [Bacillus subtilis]
1091	Bt1Gc1430	Bt1G1797	5550-4782	g1518679	559	584	9.90E-57	44	100	(U60901) orf [Bacillus subtilis]
1091	Bt1Gc1430	Bt1G1798	5832-5565	g2633913	257	273	9.00E-24	58	99	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
1092	Bt1Gc1432	Bt1G1799	1176-460	g1653574	68	134	4.90E-07	27	56	(D90914) negative aliphatic amidase regulator

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat. nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description *
1092	Bt1Gc1432	Bt1G1800	2351-1497	g2982874	211	342	4.40E-31	29	100	[Synechocystis sp.] (AE000675) cobalamin synthesis related protein CobW [Aquifex aeolicus]
1092	Bt1Gc1432	Bt1G1801	3196-2464	g2619006	853	875	1.40E-87	66	55	(AF027868) putative L-amino acid oxidase precursor [Bacillus subtilis]
1093	Bt1Gc1437	Bt1G1802	165-587	g1730943	133	148	1.60E-10	28	99	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN UVRX-ILVA INTERGENIC REGION [Bacillus subtilis]
1093	Bt1Gc1437	Bt1G1803	736-1911	g1001709	880	869	6.20E-87	43	100	(D64004) hypothetical protein [Synechocystis sp.]
1093	Bt1Gc1437	Bt1G1804	3654-2830	g1170787	410	349	7.90E-32	38	100	OUTER MEMBRANE LIPOPROTEIN 2 PRECURSOR (PLP2) [Pasteurella haemolytica]
1093	Bt1Gc1437	Bt1G1805	4229-3676	g401471	364	213	2.00E-17	40	85	HYPOTHETICAL ABC TRANSPORTER PERMEASE PROTEIN YAE [Escherichia coli]
1094	Bt1Gc1436	Bt1G1806	1912-1148	g3257647	464	502	4.80E-48	37	100	(AP000005) 253aa long hypothetical ATP-binding transport protein [Pyrococcus horikoshii]
1094	Bt1Gc1436	Bt1G1807	2719-2931	g3219982	208	222	2.30E-18	61	99	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR AF1627 [Archaeoglobus fulgidus]
1094	Bt1Gc1436	Bt1G1808	4699-4229	g2851600	253	274	7.00E-24	35	70	HYPOTHETICAL 25.3 KD PROTEIN IN RIMI-PRFC INTERGENIC REGION [Escherichia coli]
1095	Bt1Gc1439	Bt1G1809	1-1181	g4103625	1155	1084	1.00E-109	57	87	(AF026470) gluconate permease [Pseudomonas aeruginosa]
1095	Bt1Gc1439	Bt1G1810	1203-2096	g1730976	999	952	1.00E-95	64	100	HYPOTHETICAL 32.8 KD PROTEIN IN NUCB-AROD INTERGENIC REGION [Bacillus subtilis]
1095	Bt1Gc1439	Bt1G1811	2845-2171	g732332	615	577	5.50E-56	49	100	HYPOTHETICAL 25.9 KD PROTEIN IN EPR-GALK INTERGENIC REGION [Bacillus subtilis]
1095	Bt1Gc1439	Bt1G1812	3099-5292	g1208461	87	113	5.90E-06	46	15	(D64004) hypothetical protein [Synechocystis sp.]
1095	Bt1Gc1439	Bt1G1813	5626-5860	g2226142	183	172	4.50E-13	51	45	(Y14079) hypothetical protein [Bacillus subtilis]
1096	Bt1Gc1441	Bt1G1814	163-861	g399775	921	921	1.90E-92	74	100	PROBABLE MENAQUINONE BIOSYNTHESIS METHYLTRANSFERASE (SPORE GERMINATION PROTEIN C2) [Bacillus subtilis]
1096	Bt1Gc1441	Bt1G1815	412-1868	g399776	1062	1056	9.50E-107	61	100	PROBABLE HEPTAPRENYL DIPHOSPHATE SYNTHASE COMPONENT II (HEPPP SYNTHASE) (SPORE GERMINATION PROTEIN C3) [Bacillus subtilis]
1096	Bt1Gc1441	Bt1G1816	1993-2439	g400405	577	577	5.50E-56	72	99	NUCLEOSIDE DIPHOSPHATE KINASE (NDK) (NDP KINASE) [Bacillus subtilis]
1096	Bt1Gc1441	Bt1G1817	2706-3343	g399058	873	765	6.50E-76	77	58	CHORISMATE SYNTHASE (5-

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1097	Bt1Gc1443	Bt1G1818	1523-974	g2634185	249	325	2.80E-29	38	99	ENOLPYRUVYL SHIKIMATE-3-PHOSPHATE PHOSPHOLYASE) (VEGETATIVE PROTEIN 216) (VEG216) [Bacillus subtilis]
1098	Bt1Gc1444	Bt1G1819	4251-4545	g2497382	424	470	1.20E-44	94	23	(Z99113) similar to thiol:disulfide interchange protein [Bacillus subtilis]
1099	Bt1Gc1442	Bt1G1820	179-1	g1730193	238	253	1.20E-21	78	17	TRANSPPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1099	Bt1Gc1442	Bt1G1821	2916-1780	g1177010	1408	1290	1.50E-131	72	79	UDP-GLUCOSE 4-EPIMERASE (GALACTOWALDENASE) (UDP-GALACTOSE 4-EPIMERASE) [Bacillus subtilis]
1099	Bt1Gc1442	Bt1G1822	3367-3049	g225559	501	531	4.10E-51	97	22	PROBABLE RNA HELICASE IN WAPA-LICT INTERGENIC REGION [Bacillus subtilis]
1100	Bt1Gc1445	Bt1G1823	794-2077	g548909	1082	1028	8.80E-104	50	100	ORF IS231C [Bacillus thuringiensis]
1100	Bt1Gc1445	Bt1G1824	2541-3027	g267523	208	227	6.70E-19	32	99	SERINE TRANSPORTER [Escherichia coli]
1100	Bt1Gc1445	Bt1G1825	3487-4081	g586554	221	309	1.40E-27	34	99	HYPOTHETICAL 18.2 KD PROTEIN IN PMI 5'REGION (ORF1) [Rhizobium meliloti]
1100	Bt1Gc1445	Bt1G1826	5444-3290	g2635881	295	304	4.60E-27	40	97	HYPOTHETICAL 21.1 KD PROTEIN IN FUS1-AGP1 INTERGENIC REGION [Saccharomyces cerevisiae]
1100	Bt1Gc1445	Bt1G1827	3535-5551	g4377189	224	275	7.40E-24	36	36	(Z99121) similar to hypothetical proteins [Bacillus subtilis]
1101	Bt1Gc1447	Bt1G1828	1127-1	g2635874	1381	1316	2.70E-134	72	48	GTP Cyclohydrtase & DHP Synthase [Chlamydia pneumoniae]
1101	Bt1Gc1447	Bt1G1829	2019-1279	g729450	1082	1082	1.70E-109	80	100	similar to hypothetical proteins [Bacillus subtilis]
1101	Bt1Gc1447	Bt1G1830	2411-2184	g2635876	262	145	3.30E-10	68	99	CARBOXYLESTERASE PRECURSOR [Bacillus stearothermophilus]
1101	Bt1Gc1447	Bt1G1831	3563-2668	g2636549	525	461	1.10E-43	41	100	(Z99121) yvaL [Bacillus subtilis]
1101	Bt1Gc1447	Bt1G1832	4058-4560	g4584100	418	449	2.00E-42	54	99	similar to hypothetical proteins [Bacillus subtilis]
1101	Bt1Gc1447	Bt1G1833	5074-4921	g3152725	151	167	9.50E-12	63	12	(AJ010132) hypothetical protein [Bacillus cereus]
1102	Bt1Gc1446	Bt1G1834	922-1	g580888	1336	1296	3.50E-132	86	53	(AF065394) enolase [Staphylococcus aureus]
1102	Bt1Gc1446	Bt1G1835	1512-952	g121635	588	563	1.70E-54	62	99	(X51477) partial dnaK gene (AA 572) [Bacillus subtilis]
1102	Bt1Gc1446	Bt1G1836	2646-1618	g3122229	1123	1143	5.80E-116	65	100	GRPE PROTEIN [Bacillus subtilis]
1102	Bt1Gc1446	Bt1G1837	3742-2780	g2104798	963	969	1.60E-97	58	100	HEAT-INDUCIBLE TRANSCRIPTION REPRESSOR HRCA [Bacillus stearothermophilus]
1103	Bt1Gc1449	Bt1G1838	736-515	g2924355	341	341	5.60E-31	95	99	(Y09446) coproporphyrinogen III oxidase [Bacillus stearothermophilus]
1103	Bt1Gc1449	Bt1G1838	736-515	g2924355	341	341	5.60E-31	95	99	(AJ223964) GerE [Bacillus subtilis]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1103	Bt1Gc1449	Bt1G1839	2170-688	g732364	570	518	9.80E-50	33	100	HYPOTHETICAL 50.0 KD PROTEIN IN UNG-ROCA INTERGENIC REGION [Bacillus subtilis]
1103	Bt1Gc1449	Bt1G1840	4251-2319	g2495515	763	832	5.20E-83	38	100	HYPOTHETICAL 50.5 KD PROTEIN IN BETT-PRPR INTERGENIC REGION [Escherichia coli]
1104	Bt1Gc1448	Bt1G1841	1571-2434	g585208	721	749	3.20E-74	49	100	GLYCEROPHOSPHORYL DIESTER PHOSPHODIESTERASE (GLYCEROPHOSPHODIESTER PHOSPHODIESTERASE) [Bacillus subtilis]
1104	Bt1Gc1448	Bt1G1842	3382-3945	g1652383	222	222	2.30E-18	30	99	(D90905) hypothetical protein [Synechocystis sp.]
1104	Bt1Gc1448	Bt1G1843	5363-3917	g1881344	1562	1592	1.50E-163	63	100	(AB001488) SIMILAR TO THE RHIZOPINE CATABOLISM (MOCR) GENE OF RHIZOBIUM MELIOTI. [Bacillus subtilis]
1105	Bt1Gc1451	Bt1G1844	568-1	g232174	556	567	6.30E-55	55	98	GLYCEROL UPTAKE OPERON ANTITERMINATOR REGULATORY PROTEIN [Bacillus subtilis]
1105	Bt1Gc1451	Bt1G1845	1112-1	g2127359	119	174	2.80E-13	35	39	hypothetical protein 3 - Clostridium perfringens [Clostridium perfringens]
1105	Bt1Gc1451	Bt1G1846	2700-1525	g2619026	1727	1729	4.60E-178	79	100	(AF027868) Yoan [Bacillus subtilis]
1105	Bt1Gc1451	Bt1G1847	4317-3123	g2226126	162	154	3.60E-11	46	99	(Y14078) Hypothetical protein [Bacillus subtilis]
1105	Bt1Gc1451	Bt1G1848	5619-3331	g126303	135	188	1.60E-14	32	63	LIPASE 3 PRECURSOR (TRIACYLGLYCEROL LIPASE) [Moraxella sp.]
1106	Bt1Gc1450	Bt1G1849	1-1051	g3914433	889	902	2.00E-90	49	73	PRISMANE PROTEIN HOMOLOG [Methanobacterium thermoautotrophicum]
1106	Bt1Gc1450	Bt1G1850	3698-4039	g3582220	208	223	1.80E-18	35	99	(AE001272) conserved hypothetical protein [Lactococcus lactis]
1106	Bt1Gc1450	Bt1G1851	4630-5294	g419694	508	544	1.70E-52	47	69	hypothetical protein 4 - Streptomyces antibioticus [Streptomyces antibioticus]
1107	Bt1Gc1452	Bt1G1852	2398-620	g2635698	556	621	1.20E-60	44	100	(Z99120) yuii [Bacillus subtilis]
1107	Bt1Gc1452	Bt1G1853	2296-1490	g1706796	779	779	2.10E-77	52	100	FERRICHRONE TRANSPORT ATP-BINDING PROTEIN FHUC [Bacillus subtilis]
1107	Bt1Gc1452	Bt1G1854	3309-1239	g1706787	942	647	2.10E-63	48	100	IRON-UPTAKE SYSTEM PROTEIN FEUC [Bacillus subtilis]
1107	Bt1Gc1452	Bt1G1855	4331-3329	g729484	1034	913	1.40E-91	58	100	IRON-UPTAKE SYSTEM PROTEIN FEUB [Bacillus subtilis]
1107	Bt1Gc1452	Bt1G1856	5251-4361	g729483	749	739	3.70E-73	50	94	IRON-UPTAKE SYSTEM BINDING PROTEIN PRECURSOR [Bacillus subtilis]
1108	Bt1Gc1438	Bt1G1857	515-1	g282367	639	672	4.70E-66	76	77	transcription initiation factor sigma H - Bacillus megaterium

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1108	Bt1Gc1438	Bt1G1858	1079-570	g586909	494	502	4.80E-48	56	99	[Bacillus megaterium] HYPOTHETICAL 19.7 KD PROTEIN IN CYSS 3'REGION [Bacillus subtilis]
1108	Bt1Gc1438	Bt1G1859	1835-1089	g586908	940	943	9.00E-95	69	100	HYPOTHETICAL TRNA/RRNA METHYLTRANSFERASE YACO [Bacillus subtilis]
1108	Bt1Gc1438	Bt1G1860	2251-1821	g2632362	374	341	5.60E-31	55	99	(Z99104) similar to hypothetical proteins [Bacillus subtilis]
1108	Bt1Gc1438	Bt1G1861	2667-2236	g549024	463	378	6.70E-35	60	31	CYSTEINYL-TRNA SYNTHETASE (CYSTEINE--TRNA LIGASE) (CYRS) [Bacillus subtilis]
1109	Bt1Gc1454	Bt1G1862	790-1	g2632007	852	740	2.90E-73	63	60	(AJ002571) Ykba [Bacillus subtilis]
1109	Bt1Gc1454	Bt1G1863	3072-2284	g267022	466	513	3.30E-49	41	100	STAGE III SPOULATION PROTEIN J PRECURSOR [Bacillus subtilis]
1110	Bt1Gc1457	Bt1G1864	1824-1471	g2073397	553	471	9.30E-45	92	99	(Y09322) membrane bound protein LytR [Bacillus cereus]
1110	Bt1Gc1457	Bt1G1865	3340-2165	g1709415	1306	985	3.20E-99	67	100	PYRIMIDINE NUCLEOSIDE TRANSPORT PROTEIN [Bacillus subtilis]
1110	Bt1Gc1457	Bt1G1866	3442-3524	g225559	109	126	3.10E-07	89	6	ORF IS231C [Bacillus thuringiensis]
1111	Bt1Gc1453	Bt1G1867	338-727	g584819	184	191	4.40E-15	39	99	ATP SYNTHASE PROTEIN I [Bacillus subtilis]
1111	Bt1Gc1453	Bt1G1868	735-1454	g114432	866	683	3.20E-67	70	100	ATP SYNTHASE A CHAIN (PROTEIN 6) [Bacillus megaterium]
1111	Bt1Gc1453	Bt1G1869	1514-1729	g114678	284	172	4.50E-13	81	99	ATP SYNTHASE C CHAIN (LIPID-BINDING PROTEIN) [thermophilic bacterium PS3]
1111	Bt1Gc1453	Bt1G1870	1850-2365	g114616	503	405	9.20E-38	61	99	ATP SYNTHASE B CHAIN [Bacillus megaterium]
1111	Bt1Gc1453	Bt1G1871	2365-2898	g1168584	444	446	4.20E-42	46	99	ATP SYNTHASE DELTA CHAIN [Bacillus stearothermophilus]
1111	Bt1Gc1453	Bt1G1872	2919-3410	g114510	773	773	9.30E-77	93	32	ATP SYNTHASE ALPHA CHAIN [Bacillus megaterium]
1112	Bt1Gc1458	Bt1G1873	817-1363	g4033455	369	404	1.20E-37	45	99	SIGNAL PEPTIDASE I P (SPASE I) (LEADER PEPTIDASE I) [Bacillus subtilis]
1113	Bt1Gc1456	Bt1G1874	1077-286	g3915949	356	343	3.40E-31	34	100	HYPOTHETICAL 28.2 KD PROTEIN IN USHA-TESA INTERGENIC REGION []
1113	Bt1Gc1456	Bt1G1875	2557-888	g2612897	226	319	1.20E-28	32	100	(AF015825) hypothetical ABC transporter [Bacillus subtilis]
1113	Bt1Gc1456	Bt1G1876	3256-2264	g2635717	1181	1154	3.90E-117	67	100	(Z99120) similar to NADH dehydrogenase [Bacillus subtilis]
1113	Bt1Gc1456	Bt1G1877	3606-3839	g2635716	288	288	2.30E-25	68	99	(Z99120) yuzB [Bacillus subtilis]
1114	Bt1Gc1460	Bt1G1878	1-1195	g2635874	1359	1335	5.20E-138	73	48	(Z99121) similar to hypothetical proteins [Bacillus subtilis]
1114	Bt1Gc1460	Bt1G1879	1331-1798	g3334327	639	591	1.80E-57	76	99	SMALL PROTEIN B HOMOLOG [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1114	Bt1Gc1460	Bt1G1880	3882-4040	g225559	223	245	3.60E-20	92	11	ORF IS231C [Bacillus thuringiensis]
1115	Bt1Gc1455	Bt1G1881	1-4368	g2982194	2021	2163	3.30E-222	35	25	(AF007865) bacitracin synthetase 1; BacA [Bacillus licheniformis]
1115	Bt1Gc1455	Bt1G1882	2982-4368	g4126805	650	669	9.70E-66	31	99	(AB016966) bacitracin synthetase 3 (BA3) [Bacillus licheniformis]
1116	Bt1Gc1461	Bt1G1883	1250-48	g729203	959	1001	6.40E-101	47	100	CYTOCHROME P450(MEG) (STEROID 15-BETA-HYDROXYLASE) (STEROID 15-BETA-MONOOXYGENASE) [Bacillus megaterium]
1116	Bt1Gc1461	Bt1G1884	1381-1706	g730644	81	111	1.30E-06	24	99	30S RIBOSOMAL PROTEIN S14 HOMOLOG [Bacillus sphaericus]
1116	Bt1Gc1461	Bt1G1885	2876-2238	g3660657	753	705	1.50E-69	65	100	(AF015628) streptogramin A acetyl transferase [Staphylococcus cohnii]
1116	Bt1Gc1461	Bt1G1886	4092-2833	g2632224	251	145	5.80E-18	23	100	(AJ222587) YkuC protein [Bacillus subtilis]
1116	Bt1Gc1461	Bt1G1887	5633-4401	g2811064	1232	1208	7.40E-123	58	100	CYTOCHROME P450 [Bacillus subtilis]
1117	Bt1Gc1462	Bt1G1888	1-354	g2633538	313	322	5.70E-29	55	84	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
1117	Bt1Gc1462	Bt1G1889	599-2528	g2633536	804	910	2.80E-91	36	82	(Z99110) similar to ATP-dependent DNA helicase [Bacillus subtilis]
1118	Bt1Gc1465	Bt1G1890	787-2154	g2226160	833	819	1.20E-81	38	100	(Y14080) hypothetical protein [Bacillus subtilis]
1118	Bt1Gc1465	Bt1G1891	2509-2255	g3688810	110	127	2.60E-08	31	99	(AF084104) hypothetical protein [Bacillus firmus]
1118	Bt1Gc1465	Bt1G1892	2967-3122	g134229	228	228	5.30E-19	87	71	SMALL, ACID-SOLUBLE SPORE PROTEIN 1 (SASP) [Sporosarcina halophila]
1119	Bt1Gc1463	Bt1G1893	1551-667	g1894751	893	877	8.90E-88	62	100	(Z92952) ywqM [Bacillus subtilis]
1119	Bt1Gc1463	Bt1G1894	1672-2637	g1172802	464	495	2.70E-47	39	100	QUINONE OXIDOREDUCTASE (NADPH:QUINONE REDUCTASE) [Pseudomonas aeruginosa]
1119	Bt1Gc1463	Bt1G1895	4785-2904	g2829611	157	241	2.20E-20	29	91	HEMK PROTEIN HOMOLOG [Synechocystis sp.]
1119	Bt1Gc1463	Bt1G1896	4643-3969	g548530	106	203	2.30E-16	25	100	PROBABLE PHOSPHOGLYCERATE MUTASE 2 (PHOSPHOGLYCEROMUTASE 2) (PGAM 2) (BPG-DEPENDENT PGAM 2) []
1119	Bt1Gc1463	Bt1G1897	5150-4228	g732075	105	174	2.80E-13	35	61	HYPOTHETICAL 20.0 KD PROTEIN IN FECI-FIMB INTERGENIC REGION (F181) [Escherichia coli]
1120	Bt1Gc1464	Bt1G1898	1861-3843	g1075867	1405	1165	2.70E-118	45	100	methy-accepting chemotaxis transducer MCPB - Bacillus subtilis []
1120	Bt1Gc1464	Bt1G1899	5278-4158	g1565245	241	320	9.40E-29	26	100	(Z80360) Unknown, highly similar to B. subtilis RapA and RapB aspartyl-phosphate phosphatases [Bacillus subtilis]
1120	Bt1Gc1464	Bt1G1900	7479-6403	g609642	942	896	8.60E-90	52	100	(L33181) CDP-D-glucose-4,6-dehydratase [Yersinia

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1121	Bt1Gc1467	Bt1G1901	228-1118	g2735505	528	565	1.00E-54	38	100	pseudotuberculosis [U96107] N5,N10-methyltetrahydromethanopterin reductase homolog [Staphylococcus carnosus]
1121	Bt1Gc1467	Bt1G1902	1716-3119	g2829796	1459	1422	1.60E-145	59	100	CYTOCHROME D UBIQUINOL OXIDASE SUBUNIT I [Bacillus subtilis]
1121	Bt1Gc1467	Bt1G1903	3106-4119	g2829797	939	907	5.90E-91	50	100	CYTOCHROME D UBIQUINOL OXIDASE SUBUNIT II [Bacillus subtilis]
1121	Bt1Gc1467	Bt1G1904	4128-4501	g2829798	201	215	1.00E-16	34	22	TRANSPORT ATP-BINDING PROTEIN CYDC [Bacillus subtilis]
1122	Bt1Gc1468	Bt1G1905	461-6	g1730924	242	283	7.80E-25	37	99	HYPOTHETICAL 17.3 KD PROTEIN IN QCRA-AROE INTERGENIC REGION [Bacillus subtilis]
1122	Bt1Gc1468	Bt1G1906	1152-616	g1730923	661	665	2.60E-65	64	99	HYPOTHETICAL 21.4 KD PROTEIN IN QCRA-AROE INTERGENIC REGION [Bacillus subtilis]
1122	Bt1Gc1468	Bt1G1907	2491-1225	g1730922	856	772	1.20E-76	42	100	HYPOTHETICAL 48.3 KD PROTEIN IN QCRA-AROE INTERGENIC REGION [Bacillus subtilis]
1122	Bt1Gc1468	Bt1G1908	3929-2853	g3123224	1126	1063	1.70E-107	59	100	HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (IMIDAZOLE ACETOL-PHOSPHATE TRANSAMINASE) [Bacillus subtilis]
1122	Bt1Gc1468	Bt1G1909	4073-4396	g225559	541	541	3.60E-52	98	22	ORF IS231C [Bacillus thuringiensis]
1123	Bt1Gc1466	Bt1G1910	942-292	g3170570	371	395	1.10E-36	39	100	(AF058302) FmE [Streptomyces roseofulvus]
1123	Bt1Gc1466	Bt1G1911	2396-1130	g1731090	216	277	3.40E-24	24	80	HYPOTHETICAL 34.6 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1124	Bt1Gc1469	Bt1G1912	2792-1133	g3688808	1790	1835	2.70E-189	60	100	(AF084104) AcsA [Bacillus firmus]
1124	Bt1Gc1469	Bt1G1913	4906-3124	g2634204	2041	2090	2.60E-216	68	100	(Z99113) similar to propionyl-CoA carboxylase [Bacillus subtilis]
1124	Bt1Gc1469	Bt1G1914	5447-4668	g2266427	747	674	2.90E-66	61	100	(Y13917) yngF [Bacillus subtilis]
1124	Bt1Gc1469	Bt1G1915	6369-5473	g2266426	847	850	6.40E-85	54	100	(Y13917) yngG [Bacillus subtilis]
1124	Bt1Gc1469	Bt1G1916	6599-6391	g4584148	106	145	7.70E-09	40	7	(AJ010111) pyruvate carboxylase [Bacillus cereus]
1124	Bt1Gc1469	Bt1G1917	6864-6645	g2266425	178	194	1.10E-14	51	16	(Y13917) yngH [Bacillus subtilis]
1125	Bt1Gc1459	Bt1G1918	957-1	g2894238	283	357	1.10E-32	30	78	(AL021841) sigJ [Mycobacterium tuberculosis]
1125	Bt1Gc1459	Bt1G1919	2278-1	g2749982	87	136	3.20E-08	44	13	(AF036705) Similar to phytoene desaturase; coded for by C. elegans cDNA CEESX74F; coded for by C. elegans cDNA yk303f4.3; coded for by C. elegans cDNA yk257d4.3; coded for by C. elegans cDNA yk303f4.5; coded for by C. elegans cDNA yk257d4.5;... [ ]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1126	Bt1Gc1470	Bt1G1921	248-1	g1730269	223	228	5.30E-19	52	73	HYPOTHETICAL 11.9 KD PROTEIN IN HMP 3'REGION [Bacillus subtilis]
1126	Bt1Gc1470	Bt1G1922	930-1	g2984362	57	118	2.40E-07	32	48	(AE000776) transcriptional regulator (TetR/AcrR family) [Aquifex aeolicus]
1126	Bt1Gc1470	Bt1G1923	1973-1000	g2633758	752	784	6.30E-78	48	100	(Z99111) similar to transcriptional regulator (LacI family) [Bacillus subtilis]
1126	Bt1Gc1470	Bt1G1924	3242-2865	g1731041	403	411	2.10E-38	60	99	HYPOTHETICAL 14.6 KD PROTEIN IN GCVT-SPOIIIAA INTERGENIC REGION [Bacillus subtilis]
1126	Bt1Gc1470	Bt1G1925	3419-4252	g2634887	1009	1009	9.10E-102	67	100	(Z99116) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1127	Bt1Gc1471	Bt1G1926	1-1155	g2293179	1119	502	1.60E-96	55	96	(AF008220) YttB [Bacillus subtilis]
1127	Bt1Gc1471	Bt1G1927	1625-2286	g3123286	1075	1027	1.10E-103	87	27	LEUCYL-TRNA SYNTHETASE (LEUCINE--TRNA LIGASE) (LEURS) [Bacillus subtilis]
1128	Bt1Gc1475	Bt1G1928	625-1	g548909	479	388	5.80E-36	46	49	SERINE TRANSPORTER [Escherichia coli]
1128	Bt1Gc1475	Bt1G1929	1842-937	g2337815	838	707	9.20E-70	55	100	(Y13937) putative YhaP protein [Bacillus subtilis]
1128	Bt1Gc1475	Bt1G1930	2523-1861	g2337814	359	390	3.60E-36	39	100	(Y13937) putative YhaQ protein [Bacillus subtilis]
1128	Bt1Gc1475	Bt1G1931	5531-2937	g3861083	85	182	5.80E-12	28	49	(AJ235272) unknown [Rickettsia prowazekii]
1129	Bt1Gc1474	Bt1G1932	519-1420	g4104606	503	473	5.70E-45	39	76	(AF036967) putative histidine kinase [Lactobacillus sakei]
1129	Bt1Gc1474	Bt1G1933	1446-3762	g4887207	231	370	4.70E-34	31	73	(AF147448) penicillin-binding protein 5 [Pseudomonas aeruginosa]
1129	Bt1Gc1474	Bt1G1934	3591-2615	g1934836	293	377	8.50E-35	44	99	(Z93940) unknown [Bacillus subtilis]
1129	Bt1Gc1474	Bt1G1935	4794-4231	g2633543	74	145	3.30E-10	23	99	(Z99110) similar to ribosomal-protein-alanine N-acetyltransferase [Bacillus subtilis]
1130	Bt1Gc1478	Bt1G1936	856-1	g116465	1017	975	3.70E-98	70	76	CITRATE SYNTHASE [Bacillus coagulans]
1130	Bt1Gc1478	Bt1G1937	1545-1261	g1405458	251	253	1.20E-21	48	99	(Z73234) YneR [Bacillus subtilis]
1130	Bt1Gc1478	Bt1G1938	3563-1525	g1652918	412	300	1.20E-26	28	100	(D90909) quinolone resistance protein NorA [Synechocystis sp.]
1130	Bt1Gc1478	Bt1G1939	5805-2078	g730275	1388	1462	9.00E-150	43	83	PENICILLIN-BINDING PROTEINS 1A/1B (PBP1) [Bacillus subtilis]
1131	Bt1Gc1472	Bt1G1940	705-220	g2635218	622	632	8.10E-62	75	99	(Z99118) similar to hypothetical proteins [Bacillus subtilis]
1131	Bt1Gc1472	Bt1G1941	999-1	g3256441	117	161	6.60E-12	31	60	(AP000001) 230aa long hypothetical HESA protein [Pyrococcus horikoshii]
1131	Bt1Gc1472	Bt1G1942	3258-1486	g3122885	2183	2133	7.10E-221	71	100	ASPARTYL-TRNA SYNTHETASE (ASPARTATE--TRNA LIGASE) (ASPRS) [Bacillus subtilis]
1131	Bt1Gc1472	Bt1G1943	4542-3270	g3122900	1517	1544	1.80E-158	68	100	HISTIDYL-TRNA SYNTHETASE (HISTIDINE--TRNA

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1132	Bt1Gc1473	Bt1G1944	3718-2480	g3688826	618	615	5.10E-60	31	100	LIGASE (HISRS) [Bacillus subtilis] (AF084104) NatB [Bacillus firmus]
1132	Bt1Gc1473	Bt1G1945	4607-3714	g2226131	930	931	1.70E-93	59	100	(Y14078) Hypothetical protein [Bacillus subtilis]
1132	Bt1Gc1473	Bt1G1946	5542-5841	g420808	408	351	4.90E-32	74	99	hypothetical 12K protein - Thermus aquaticus [Thermus aquaticus]
1133	Bt1Gc1476	Bt1G1947	2433-276	g2635804	75	140	6.10E-09	35	18	(Z99120) gerAC [Bacillus subtilis]
1133	Bt1Gc1476	Bt1G1948	3503-1582	g1731309	302	401	2.40E-37	39	64	HYPOTHETICAL 37.2 KD PROTEIN IN IDH-DEOR INTERGENIC REGION [Bacillus subtilis]
1133	Bt1Gc1476	Bt1G1949	3817-2421	g1652140	1320	1369	6.50E-140	54	100	(D90903) glutamate decarboxylase [Synechocystis sp.]
1133	Bt1Gc1476	Bt1G1950	4103-5909	g1881234	199	222	2.30E-18	43	48	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
1133	Bt1Gc1476	Bt1G1951	5763-4738	g586030	1547	1547	8.90E-159	83	100	SPORE PHOTOPRODUCT LYASE [Bacillus subtilis]
1133	Bt1Gc1476	Bt1G1952	6026-4799	g421525	157	217	7.70E-18	36	99	orfY - Bacillus subtilis (fragment) []
1134	Bt1Gc1479	Bt1G1953	199-2262	g124464	3426	3205	0	93	100	IMMUNE INHIBITOR A PRECURSOR [Bacillus thuringiensis]
1134	Bt1Gc1479	Bt1G1954	2431-3801	g732362	1417	1317	2.10E-134	56	100	PROBABLE ALDEHYDE DEHYDROGENASE YWDH [Bacillus subtilis]
1134	Bt1Gc1479	Bt1G1955	4027-4763	g3341854	730	783	8.10E-78	60	66	(AF077856) PotA [Actinobacillus actinomycetemcomitans]
1134	Bt1Gc1479	Bt1G1956	5009-5751	g3341855	564	516	1.60E-49	44	86	(AF077856) PotB [Actinobacillus actinomycetemcomitans]
1135	Bt1Gc1477	Bt1G1957	136-1	g131631	125	143	7.50E-10	61	19	PHOSPHORIBOSYLAMINOIMIDAZOLE-SUCCINOCARBOXAMIDE SYNTHASE (SAICAR SYNTHETASE) (VEGETATIVE PROTEIN 286A) (VEG286A) [Bacillus subtilis]
1135	Bt1Gc1477	Bt1G1958	1526-228	g131635	1883	1895	1.20E-195	84	100	ADENYLOSUCCINATE LYASE (ADENYLOSUCCINASE) (ASL) [Bacillus subtilis]
1135	Bt1Gc1477	Bt1G1959	2677-1538	g131644	1095	1124	5.90E-114	57	100	PHOSPHORIBOSYLAMINOIMIDAZOLE CARBOXYLASE ATPASE SUBUNIT (AIR CARBOXYLASE) (AIRC) [Bacillus subtilis]
1135	Bt1Gc1477	Bt1G1960	3156-2671	g131626	642	647	2.10E-63	78	99	PHOSPHORIBOSYLAMINOIMIDAZOLE CARBOXYLASE CATALYTIC SUBUNIT (AIR CARBOXYLASE) (AIRC) [Bacillus subtilis]
1136	Bt1Gc1480	Bt1G1961	1894-30	g1894770	105	168	2.10E-12	30	43	(Z92954) product similar to Bacillus subtilis YxeH and YcsE proteins and to E. coli YidA protein [Bacillus subtilis]
1136	Bt1Gc1480	Bt1G1962	2425-947	g1304006	1493	1415	8.60E-145	59	100	(D84432) SpoVAF [Bacillus subtilis]
1136	Bt1Gc1480	Bt1G1963	3337-2378	g730786	1001	1060	3.60E-107	61	100	STAGE V SPORULATION PROTEIN AE [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1136	Bt1Gc1480	Bt1G1964	4357-3341	g730785	1215	1121	1.20E-113	68	100	STAGE V SPORULATION PROTEIN AD [Bacillus subtilis]
1136	Bt1Gc1480	Bt1G1965	4646-4374	g730784	375	375	1.40E-34	79	60	STAGE V SPORULATION PROTEIN AC [Bacillus subtilis]
1137	Bt1Gc1482	Bt1G1966	479-1	g3256494	82	159	1.10E-11	31	73	(AP000001) 216aa long hypothetical alanyl-tRNA synthetase [Pyrococcus horikoshii]
1137	Bt1Gc1482	Bt1G1967	1364-948	g600725	64	159	1.60E-10	32	19	(U09549) flagellar hook FlgE [Helicobacter pylori]
1137	Bt1Gc1482	Bt1G1968	2917-1413	g544314	231	307	9.10E-32	26	100	FLAGELLAR BIOSYNTHESIS PROTEIN FLHF (FLAGELLA ASSOCIATED GTP-BINDING PROTEIN) [Bacillus subtilis]
1137	Bt1Gc1482	Bt1G1969	4792-2738	g544312	1048	1067	6.50E-108	37	100	FLAGELLAR BIOSYNTHESIS PROTEIN FLHA [Bacillus subtilis]
1137	Bt1Gc1482	Bt1G1970	5898-4831	g544313	542	334	4.90E-52	32	100	FLAGELLAR BIOSYNTHETIC PROTEIN FLHB [Bacillus subtilis]
1137	Bt1Gc1482	Bt1G1971	6232-5888	g2984202	86	120	4.20E-07	20	44	(AE000765) flagellar biosynthetic protein FlhR [Aquifex aeolicus]
1138	Bt1Gc1485	Bt1G1972	1784-1251	g1750115	93	121	1.10E-07	26	99	(U66480) YnaD [Bacillus subtilis]
1138	Bt1Gc1485	Bt1G1973	3686-2219	g2352096	81	198	7.90E-16	28	100	(U97022) orf; similar to serine/threonine protein phosphatase [Fervidobacterium islandicum]
1138	Bt1Gc1485	Bt1G1974	4303-3881	g2226205	191	228	5.30E-19	34	99	(Y14082) hypothetical protein [Bacillus subtilis]
1138	Bt1Gc1485	Bt1G1975	4430-5825	g1881332	986	970	1.20E-97	45	100	(AB001488) SIMILAR TO THE RHIZOPINE CATABOLISM (MOCR) GENE OF RHIZOBIUM MELIOTI. [Bacillus subtilis]
1138	Bt1Gc1485	Bt1G1976	6705-5951	g1731008	308	258	3.50E-22	30	85	HYPOTHETICAL 32.5 KD PROTEIN IN CCCA-SODA INTERGENIC REGION [Bacillus subtilis]
1139	Bt1Gc1484	Bt1G1977	120-1	g1731085	114	114	6.30E-07	50	25	HYPOTHETICAL 17.9 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1139	Bt1Gc1484	Bt1G1978	2030-1155	g2650107	309	354	2.30E-32	27	100	(AE001068) conserved hypothetical protein [Archaeoglobus fulgidus]
1139	Bt1Gc1484	Bt1G1979	2158-3600	g1881344	759	855	1.90E-85	36	100	(AB001488) SIMILAR TO THE RHIZOPINE CATABOLISM (MOCR) GENE OF RHIZOBIUM MELIOTI. [Bacillus subtilis]
1139	Bt1Gc1484	Bt1G1980	3819-4271	g1731102	263	291	1.10E-25	42	99	HYPOTHETICAL 17.1 KD PROTEIN IN DNAG/DNAE 5'REGION (P17) [Listeria monocytogenes]
1139	Bt1Gc1484	Bt1G1981	4665-4360	g2636123	303	317	1.90E-28	65	99	(Z99122) ywsA [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1140	Bt1Gc1481	Bt1G1982	805-74	g3425862	661	602	1.20E-58	55	100	(AJ007731) 3-ketoacyl-ACP/CoA reductase [Streptomyces coelicolor]
1140	Bt1Gc1481	Bt1G1983	1770-2642	g3122275	849	868	8.00E-87	58	100	PUTATIVE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (TRANSAMINASE B) (BCAT) [Archaeoglobus fulgidus]
1140	Bt1Gc1481	Bt1G1984	2680-4392	g1770066	1620	1666	2.20E-171	55	100	(Z75208) acetolactate synthase large subunit [Bacillus subtilis]
1140	Bt1Gc1481	Bt1G1985	4640-5667	g585314	1099	1121	1.20E-113	61	100	KETOL-ACID REDUCTOISOMERASE (ACETOHYDROXY-ACID ISOMEROREDUCTASE) (ALPHA-KETO-BETA-HYDROXYLACIL REDUCTOISOMERASE) [Bacillus subtilis]
1141	Bt1Gc1487	Bt1G1986	1395-34	g585057	1823	1764	9.00E-182	78	100	REPLICATIVE DNA HELICASE [Bacillus subtilis]
1141	Bt1Gc1487	Bt1G1987	1869-1423	g133029	569	263	3.50E-32	75	99	50S RIBOSOMAL PROTEIN L9 (BL17) []
1141	Bt1Gc1487	Bt1G1988	3839-1869	g586814	2094	1973	6.40E-204	62	100	HYPOTHETICAL 74.3 KD PROTEIN IN RPLI-COTF INTERGENIC REGION [Bacillus subtilis]
1141	Bt1Gc1487	Bt1G1989	4853-3924	g586815	457	393	1.70E-36	32	100	HYPOTHETICAL 34.5 KD PROTEIN IN RPLI-COTF INTERGENIC REGION [Bacillus subtilis]
1141	Bt1Gc1487	Bt1G1990	5167-4934	g585939	353	295	4.20E-26	88	99	30S RIBOSOMAL PROTEIN S18 (BS21) []
1141	Bt1Gc1487	Bt1G1991	5725-5216	g586039	633	625	4.50E-61	73	99	SINGLE-STRAND BINDING PROTEIN (SSB) (HELIX-DESTABILIZING PROTEIN) [Bacillus subtilis]
1141	Bt1Gc1487	Bt1G1992	6042-5758	g585949	380	380	4.10E-35	73	99	30S RIBOSOMAL PROTEIN S6 (BS9) [Bacillus subtilis]
1142	Bt1Gc1486	Bt1G1993	674-1	g2633471	881	829	1.10E-82	73	54	(Z99109) similar to 3-oxoacyl- acyl-carrier protein synthase [Bacillus subtilis]
1142	Bt1Gc1486	Bt1G1994	1638-709	g2633470	1207	1167	1.60E-118	71	100	(Z99109) similar to 3-oxoacyl- acyl-carrier protein synthase [Bacillus subtilis]
1142	Bt1Gc1486	Bt1G1995	2196-2008	g2633468	162	168	1.20E-12	54	98	(Z99109) yjzA [Bacillus subtilis]
1142	Bt1Gc1486	Bt1G1996	3884-3153	g2564028	489	507	1.40E-48	43	100	(D86376) unnamed protein product [Bacillus subtilis]
1142	Bt1Gc1486	Bt1G1997	4737-4253	g3150046	364	371	4.00E-33	46	19	(AF016634) ClpB chaperone homolog [Lactococcus lactis subsp. cremoris]
1143	Bt1Gc1483	Bt1G1998	621-2484	g1772644	217	293	6.80E-26	41	61	(U62055) orfR gene product [Bacillus subtilis]
1143	Bt1Gc1483	Bt1G1999	2471-1478	g2190588	316	251	1.50E-20	31	57	(U75364) putative ABC transporter subunit [Rhodopseudomonas palustris]
1143	Bt1Gc1483	Bt1G2000	3361-2038	g3644001	423	244	5.70E-33	30	74	(AF087482) putative membrane spanning protein [Pseudomonas aeruginosa]
1143	Bt1Gc1483	Bt1G2001	4813-3364	g2649779	413	453	7.50E-43	39	100	(AE001047) branched-chain amino acid ABC transporter,



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1144	Bt1Gc1488	Bt1G2002	796-1	g2842762	461	528	1.40E-49	45	22	ATP-binding protein (braF-2) [Archaeoglobus fulgidus]
1144	Bt1Gc1488	Bt1G2003	2451-735	g2145402	1433	1456	3.90E-149	50	93	5-METHYLTETRAHYDROFOLATE--HOMOCYSTEINE METHYLTRANSFERASE (METHIONINE SYNTHASE, VITAMIN-B12 DEPENDENT) (MS) [Homo sapiens]
1145	Bt1Gc1489	Bt1G2004	1742-75	g2501186	1854	1685	2.10E-173	63	100	FORMATE--TETRAHYDROFOLATE LIGASE (FORMYLTETRAHYDROFOLATE SYNTHETASE) (FHS) (FTHFS) [Streptococcus mutans]
1145	Bt1Gc1489	Bt1G2005	2535-915	g2650105	80	164	3.20E-12	24	74	(AE001068) chloroplast inner envelope membrane protein [Archaeoglobus fulgidus]
1145	Bt1Gc1489	Bt1G2006	5259-4828	g2293237	399	377	8.50E-35	54	99	(AF008220) YtfJ [Bacillus subtilis]
1146	Bt1Gc1495	Bt1G2007	1-2246	g2496166	84	243	1.20E-19	24	52	HYPOTHETICAL PROTEIN MJ1236 [Methanococcus jannaschii]
1146	Bt1Gc1495	Bt1G2008	2453-2037	g2116755	334	286	3.80E-25	43	99	(D86418) YfmQ [Bacillus subtilis]
1146	Bt1Gc1495	Bt1G2009	6156-5242	g1731074	113	206	1.10E-16	23	100	HYPOTHETICAL 28.2 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1147	Bt1Gc1494	Bt1G2010	3374-4288	g3080402	348	236	4.60E-38	34	100	(AL022603) putative NADPH quinone oxidoreductase [Arabidopsis thaliana]
1148	Bt1Gc1492	Bt1G2011	567-1	g1731127	708	717	8.00E-71	72	63	HYPOTHETICAL 32.9 KD PROTEIN IN BLTR-SPOIIC INTERGENIC REGION [Bacillus subtilis]
1148	Bt1Gc1492	Bt1G2012	1400-849	g1731125	416	445	5.30E-42	44	99	HYPOTHETICAL 20.7 KD PROTEIN IN BLTR-SPOIIC INTERGENIC REGION [Bacillus subtilis]
1148	Bt1Gc1492	Bt1G2013	1686-1495	g1731123	209	218	6.00E-18	67	98	HYPOTHETICAL 7.0 KD PROTEIN IN BLTR-SPOIIC INTERGENIC REGION [Bacillus subtilis]
1148	Bt1Gc1492	Bt1G2014	2323-1950	g1731041	133	173	3.50E-13	33	99	HYPOTHETICAL 14.6 KD PROTEIN IN GCVT-SPOIIIA INTERGENIC REGION [Bacillus subtilis]
1148	Bt1Gc1492	Bt1G2015	3602-2265	g2792490	772	730	3.30E-72	39	100	(AF041467) coenzyme A disulfide reductase [Staphylococcus aureus]
1148	Bt1Gc1492	Bt1G2016	5271-4309	g730056	598	696	1.30E-68	45	100	EXTRACELLULAR METALLOPROTEASE PRECURSOR [Bacillus subtilis]
1149	Bt1Gc1493	Bt1G2017	1224-2246	g4753135	64	168	5.30E-12	35	29	(U52844) putative glycosyltransferase [Serratia marcescens]
1149	Bt1Gc1493	Bt1G2018	1530-1838	g1495283	91	160	1.50E-11	42	35	(Z71928) hypothetical protein [Bacillus subtilis]
1149	Bt1Gc1493	Bt1G2019	4192-3269	g2494668	418	449	2.00E-42	33	100	PUTATIVE UDP-GLUCOSE 4-EPIMERASE (GALACTOWALDENASE) (UDP-GALACTOSE 4-EPIMERASE) [Methanococcus jannaschii]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1149	Bt1Gc1493	Bt1G2020	5034-4644	g2497392	668	670	7.60E-66	99	52	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
1150	Bt1Gc1496	Bt1G2021	41-778	g1805478	656	659	1.10E-64	52	100	(D50453) homologue of pectin degradation repressor KdGR of <i>Erwinia chrysanthemi</i> [Bacillus subtilis]
1150	Bt1Gc1496	Bt1G2022	799-2508	g1175716	944	1000	8.20E-101	39	100	HYPOTHETICAL 63.8 KD PROTEIN IN SIPU-PBPC INTERGENIC REGION [Bacillus subtilis]
1150	Bt1Gc1496	Bt1G2023	2521-3279	g2290993	684	697	1.10E-68	53	100	(AF006000) unknown [Bordetella pertussis]
1150	Bt1Gc1496	Bt1G2024	4981-5628	g1172045	612	639	1.50E-62	55	100	PYRROLIDONE-CARBOXYLATE PEPTIDASE (5-OXOPROLYL-PEPTIDASE) (PYROGLUTAMYL-PEPTIDASE I) [Bacillus amyloliquefaciens]
1150	Bt1Gc1496	Bt1G2025	5716-6432	g2226221	531	555	1.20E-53	46	100	(Y14082) hypothetical protein [Bacillus subtilis]
1151	Bt1Gc1491	Bt1G2026	1958-1251	g141438	355	378	6.70E-35	33	100	HYPOTHETICAL 27.3 KD PROTEIN IN TYRZ-SACY INTERGENIC REGION (ORF1) [Bacillus subtilis]
1151	Bt1Gc1491	Bt1G2027	3267-3956	g2633807	738	751	2.00E-74	62	100	(Z99111) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis]
1152	Bt1Gc1490	Bt1G2028	84-1906	g2618848	781	714	1.70E-70	40	100	(AF017113) YvmA [Bacillus subtilis]
1152	Bt1Gc1490	Bt1G2029	2725-1305	g3122316	772	871	3.80E-87	45	100	PUTATIVE SENSOR PROTEIN KDPD [Synechocystis sp.]
1152	Bt1Gc1490	Bt1G2030	3076-2492	g2275251	444	483	5.00E-46	47	100	(U44892) KdpC [Clostridium acetobutylicum]
1152	Bt1Gc1490	Bt1G2031	4450-3094	g3121784	1537	1382	2.70E-141	66	66	POTASSIUM-TRANSPORTING ATPASE B CHAIN (ATP PHOSPHOHYDROLASE) [Clostridium acetobutylicum]
1153	Bt1Gc1499	Bt1G2032	1229-591	g1303911	335	250	2.40E-21	41	100	(D84432) SpoIIAH [Bacillus subtilis]
1153	Bt1Gc1499	Bt1G2033	1898-1233	g1710826	426	418	3.90E-39	42	100	STAGE III SPORULATION PROTEIN AG [Bacillus subtilis]
1153	Bt1Gc1499	Bt1G2034	2529-1903	g1710825	300	272	1.10E-23	35	100	STAGE III SPORULATION PROTEIN AF [Bacillus subtilis]
1153	Bt1Gc1499	Bt1G2035	3711-2545	g1000363	1204	423	7.80E-74	59	100	(U35252) SpoIIAE [Bacillus subtilis]
1153	Bt1Gc1499	Bt1G2036	4123-3725	g1710823	447	377	8.50E-35	66	99	STAGE III SPORULATION PROTEIN AD [Bacillus subtilis]
1153	Bt1Gc1499	Bt1G2037	4335-4132	g1710822	208	212	2.60E-17	56	99	STAGE III SPORULATION PROTEIN AC [Bacillus subtilis]
1153	Bt1Gc1499	Bt1G2038	4865-4353	g1710821	500	452	9.60E-43	54	99	STAGE III SPORULATION PROTEIN AB [Bacillus subtilis]
1153	Bt1Gc1499	Bt1G2039	5784-4865	g730774	939	971	9.70E-98	61	100	STAGE III SPORULATION PROTEIN AA [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1153	Bt1Gc1499	Bt1G2040	6270-6921	g141186	91	152	5.90E-11	25	98	HYPOTHETICAL P20 PROTEIN [Bacillus licheniformis]
1154	Bt1Gc1497	Bt1G2041	2352-1510	g2661707	316	416	6.30E-39	35	100	(AL009204) thiosulfate sulfurtransferase [Streptomyces coelicolor]
1154	Bt1Gc1497	Bt1G2042	2624-2370	g1730892	217	219	4.70E-18	51	99	HYPOTHETICAL 9.9 KD PROTEIN IN BCSEA-DEGR INTERGENIC REGION [Bacillus subtilis]
1154	Bt1Gc1497	Bt1G2043	3937-2657	g2635902	401	431	1.60E-40	54	99	(Z99121) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1154	Bt1Gc1497	Bt1G2044	5658-2476	g1730891	603	733	9.60E-72	29	59	HYPOTHETICAL 137.4 KD PROTEIN IN BCSEA-DEGR INTERGENIC REGION [Bacillus subtilis]
1155	Bt1Gc1500	Bt1G2045	1183-71	g2293327	794	628	2.20E-61	43	100	(AF008220) YtwI [Bacillus subtilis]
1155	Bt1Gc1500	Bt1G2046	1746-2207	g2293266	522	410	2.70E-38	68	99	(AF008220) YtwI [Bacillus subtilis]
1155	Bt1Gc1500	Bt1G2047	2628-3743	g3123196	1413	1231	2.70E-125	73	100	CITRATE SYNTHASE II [Bacillus subtilis]
1155	Bt1Gc1500	Bt1G2048	3898-5172	g2168132	1917	1794	6.00E-185	86	100	(Y13358) isocitrate dehydrogenase [Bacillus israeli]
1155	Bt1Gc1500	Bt1G2049	5217-6152	g2497856	1362	1229	4.40E-125	85	100	MALATE DEHYDROGENASE [Bacillus israeli]
1156	Bt1Gc1501	Bt1G2050	1231-1894	g586824	421	462	8.40E-44	45	94	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN IN COTF-TETB INTERGENIC REGION [Bacillus subtilis]
1156	Bt1Gc1501	Bt1G2051	2932-4098	g134396	1585	1585	8.40E-163	78	46	PREPROTEIN TRANSLOCASE SECA SUBUNIT [Bacillus subtilis]
1157	Bt1Gc1498	Bt1G2052	2409-1474	g2619048	155	257	4.40E-22	24	100	(AF027868) transcription regulator [Bacillus subtilis]
1157	Bt1Gc1498	Bt1G2053	2521-3039	g1176952	460	441	1.40E-41	50	99	HYPOTHETICAL 19.6 KD PROTEIN IN ACDA 5'REGION [Bacillus subtilis]
1157	Bt1Gc1498	Bt1G2054	4247-3080	g3687664	869	707	9.20E-70	44	84	(AF049873) sensor protein [Lactococcus lactis]
1158	Bt1Gc1502	Bt1G2055	1-346	g3122814	514	462	8.40E-44	93	30	30S RIBOSOMAL PROTEIN S1 HOMOLOG [Bacillus cereus]
1158	Bt1Gc1502	Bt1G2056	773-1408	g1730909	654	654	3.80E-64	60	100	HYPOTHETICAL 22.6 KD PROTEIN IN CMK-GPSA INTERGENIC REGION [Bacillus subtilis]
1158	Bt1Gc1502	Bt1G2057	2574-3467	g1730914	927	828	1.40E-82	61	100	HYPOTHETICAL 32.9 KD PROTEIN IN CMK-GPSA INTERGENIC REGION [Bacillus subtilis]
1158	Bt1Gc1502	Bt1G2058	3774-5081	g1730915	1838	1719	5.30E-177	79	100	HYPOTHETICAL 48.8 KD GTP-BINDING PROTEIN IN CMK-GPSA INTERGENIC REGION [Bacillus subtilis]
1158	Bt1Gc1502	Bt1G2059	5103-6141	g1169997	1154	1173	3.80E-119	64	100	GLYCEROL-3-PHOSPHATE DEHYDROGENASE [Bacillus subtilis]
1158	Bt1Gc1502	Bt1G2060	6193-6492	g2633535	195	221	2.90E-18	45	99	(Z99110) yjcC [Bacillus subtilis]
1158	Bt1Gc1502	Bt1G2061	6759-6959	g1730916	191	134	4.80E-09	55	99	HYPOTHETICAL 7.5 KD PROTEIN IN GPSA-SPOIVA

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1159	Bt1Gc1503	Bt1G2062	1707-1249	g1881322	633	562	2.10E-54	82	99	INTERGENIC REGION [Bacillus subtilis] (AB001488) SIMILAR TO TRANSCRIPTION FACTOR OF MYXOCOCCUS XANTHUS. [Bacillus subtilis]
1159	Bt1Gc1503	Bt1G2063	2942-1980	g1170422	776	821	7.60E-82	49	100	FORMIMINOGLUTAMASE (FORMIMINOGLUTAMATE HYDROLASE) [Bacillus subtilis]
1159	Bt1Gc1503	Bt1G2064	4198-2933	g1170424	1334	1319	1.30E-134	62	100	IMIDAZOLONEPROPIONASE (IMIDAZOLONE-5-PROPIONATE HYDROLASE) [Bacillus subtilis]
1159	Bt1Gc1503	Bt1G2065	5116-4205	g1170427	1196	1076	7.20E-109	72	55	UROCANATE HYDRATASE (UROCANASE) (IMIDAZOLONEPROPIONATE HYDROLASE) [Bacillus subtilis]
1160	Bt1Gc1504	Bt1G2066	183-2148	g3329623	47	138	3.00E-07	23	69	(AF078790) No definition line found [Caenorhabditis elegans]
1160	Bt1Gc1504	Bt1G2067	1329-907	g1176526	381	410	2.70E-38	54	99	HYPOTHETICAL 16.2 KD PROTEIN IN CHEV 3'REGION (ORFB) []
1160	Bt1Gc1504	Bt1G2068	1525-4004	g2313368	286	457	2.80E-43	26	90	(AE000546) guanosine pentaphosphate phosphohydrolase (gppA) [Helicobacter pylori 26695]
1160	Bt1Gc1504	Bt1G2069	3154-5287	g1001781	1418	1402	2.10E-143	46	100	(D64005) polyphosphate kinase [Synecocystis sp.]
1161	Bt1Gc1505	Bt1G2070	951-1	g3913544	933	883	2.00E-88	58	82	1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE (DXP REDUCTOISOMERASE) [Bacillus subtilis]
1161	Bt1Gc1505	Bt1G2071	1344-2339	g1711656	348	440	1.80E-41	35	100	TRYPTOPHANYL-TRNA SYNTHETASE (TRYPTOPHAN--TRNA LIGASE) (TRPRS) [Escherichia coli]
1161	Bt1Gc1505	Bt1G2072	4571-2333	g1177029	104	186	7.10E-14	29	37	HYPOTHETICAL 45.4 KD PROTEIN IN THIAMINASE 1 5'REGION []
1161	Bt1Gc1505	Bt1G2073	5646-5959	g4539176	68	129	5.50E-08	34	37	(AL049485) hypothetical protein [Streptomyces coelicolor]
1162	Bt1Gc1507	Bt1G2074	1-424	g1176978	233	262	1.30E-22	42	89	HYPOTHETICAL 18.1 KD PROTEIN IN GNTR-HTPG INTERGENIC REGION [Bacillus subtilis]
1162	Bt1Gc1507	Bt1G2075	527-2951	g2495406	266	325	2.80E-29	31	100	SUCROSE OPERON REPRESSOR (SCR OPERON REGULATORY PROTEIN) [Streptococcus mutans]
1162	Bt1Gc1507	Bt1G2076	2737-1421	g2635701	1411	1137	2.50E-115	63	100	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1162	Bt1Gc1507	Bt1G2077	3819-2881	g1644225	510	522	3.70E-50	36	100	(D64127) unknown [Bacillus subtilis]
1162	Bt1Gc1507	Bt1G2078	6177-3207	g1724012	1021	1190	6.00E-121	44	45	HYPOTHETICAL 132.7 KD PROTEIN IN CSPB-GLPP INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1163	Bt1Gc1508	Bt1G2079	807-1	g3334447	798	618	2.50E-60	61	62	HYPOTHETICAL 44.9 KD PROTEIN IN HOM-MRGA INTERGENIC REGION [Bacillus subtilis]
1163	Bt1Gc1508	Bt1G2080	1398-817	g1175028	622	573	1.50E-55	66	99	XANTHINE PHOSPHORIBOSYLTRANSFERASE [Bacillus subtilis]
1163	Bt1Gc1508	Bt1G2081	2471-1728	g4512426	660	661	6.90E-65	48	100	(AB017508) map homologue (identity of 81% to B. subtilis) [Bacillus halodurans]
1163	Bt1Gc1508	Bt1G2082	4661-3153	g1730961	1663	1676	1.90E-172	61	100	HYPOTHETICAL 58.2 KD PROTEIN IN KDG-T-XPT INTERGENIC REGION [Bacillus subtilis]
1163	Bt1Gc1508	Bt1G2083	6703-4772	g1730960	2000	1950	1.80E-201	61	100	PROBABLE ATP-DEPENDENT HELICASE IN COTD-KDUD INTERGENIC REGION [Bacillus subtilis]
1164	Bt1Gc1506	Bt1G2084	2396-1499	g3261756	416	546	1.10E-52	41	47	(Z95209) hypothetical protein Rv0939 [Mycobacterium tuberculosis]
1164	Bt1Gc1506	Bt1G2085	3581-2466	g3334227	954	964	5.30E-97	55	100	4-HYDROXYPHENYLPYRUVATE DIOXYGENASE (4HPPD) (HPD) [Streptomyces avermitilis]
1165	Bt1Gc1510	Bt1G2086	1341-325	g1730193	1489	1495	2.90E-153	80	100	UDP-GLUCOSE 4-EPIMERASE (GALACTOWALDENASE) (UDP-GALACTOSE 4-EPIMERASE) [Bacillus subtilis]
1165	Bt1Gc1510	Bt1G2087	2813-1494	g1894744	1452	1460	1.50E-149	63	100	(Z92952) product similar to Pseudomonas aeruginosa GDP-mannose 6-dehydrogenase protein [Bacillus subtilis]
1165	Bt1Gc1510	Bt1G2088	3640-2885	g2454562	748	755	7.50E-75	55	100	(AF015609) unknown [Bacillus subtilis]
1165	Bt1Gc1510	Bt1G2089	5239-4559	g2650003	123	232	2.60E-19	30	62	(AE001062) LPS biosynthesis protein, putative [Archaeoglobus fulgidus]
1166	Bt1Gc1512	Bt1G2090	254-1	g133188	232	249	3.10E-21	62	34	RIBONUCLEASE PH (RNASE PH) (TRNA NUCLEOTIDYLTRANSFERASE) [Bacillus subtilis]
1166	Bt1Gc1512	Bt1G2091	1440-389	g2507032	874	930	2.10E-93	52	100	GERMINATION PROTEIN GERM [Bacillus subtilis]
1166	Bt1Gc1512	Bt1G2092	2426-1620	g2462086	1342	1286	4.00E-131	97	100	(Y09719) putative glutamate racemase protein [Bacillus cereus]
1166	Bt1Gc1512	Bt1G2094	3153-2717	g1351709	151	210	4.20E-17	32	69	HYPOTHETICAL 24.2 KD PROTEIN C11D3.13 IN CHROMOSOME I [Schizosaccharomyces pombe]
1167	Bt1Gc1509	Bt1G2095	400-1	g137192	577	510	6.90E-49	82	22	EXCINUCLEASE ABC SUBUNIT C [Bacillus subtilis]
1167	Bt1Gc1509	Bt1G2096	864-553	g135765	425	373	2.30E-34	75	99	THIOREDOXIN (TRX) [Bacillus subtilis]
1167	Bt1Gc1509	Bt1G2097	2052-1079	g2494399	1253	1178	1.10E-119	73	100	ELECTRON TRANSFER FLAVOPROTEIN ALPHA-SUBUNIT (ALPHA-ETF) (ELECTRON TRANSFER FLAVOPROTEIN LARGE SUBUNIT) (ETFLS) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1167	Bt1Gc1509	Bt1G2098	2862-2092	g2494401	977	837	1.50E-83	74	100	ELECTRON TRANSFER FLAVOPROTEIN BETA-SUBUNIT (BETA-ETF) (ELECTRON TRANSFER FLAVOPROTEIN SMALL SUBUNIT) (ETFSS) [Bacillus subtilis]
1167	Bt1Gc1509	Bt1G2099	3715-2944	g1770040	648	556	9.20E-54	52	100	(Z75208) hypothetical protein [Bacillus subtilis]
1167	Bt1Gc1509	Bt1G2100	4317-3736	g1770039	625	628	2.20E-61	65	99	(Z75208) hypothetical protein [Bacillus subtilis]
1167	Bt1Gc1509	Bt1G2101	6115-4436	g2829810	1805	1814	4.50E-187	60	100	LONG-CHAIN-FATTY-ACID--COA LIGASE (LONG-CHAIN ACYL-COA SYNTHETASE) [Bacillus subtilis]
1168	Bt1Gc1513	Bt1G2102	3615-2749	g2497513	642	643	5.50E-63	45	100	HOMOSERINE KINASE (HK) [Streptococcus pneumoniae]
1168	Bt1Gc1513	Bt1G2103	4679-3624	g135811	1768	1631	1.10E-167	99	100	THREONINE SYNTHASE [Bacillus sp.]
1168	Bt1Gc1513	Bt1G2104	5964-4678	g2635723	1137	1067	6.50E-108	52	100	(Z99120) homoserine dehydrogenase [Bacillus subtilis]
1169	Bt1Gc1516	Bt1G2105	1172-1	g1724012	894	949	2.10E-95	47	33	HYPOTHETICAL 132.7 KD PROTEIN IN CSPB-GLPP INTERGENIC REGION [Bacillus subtilis]
1169	Bt1Gc1516	Bt1G2106	2503-1673	g3914460	721	718	6.30E-71	52	100	PUTATIVE NON-HEME CHLOROPEROXIDASE (CHLORIDE PEROXIDASE) [Synechocystis sp.]
1169	Bt1Gc1516	Bt1G2107	2997-3623	g140338	669	684	2.50E-67	63	100	HYPOTHETICAL 23.1 KD PROTEIN IN DMSC-PFLA INTERGENIC REGION [Escherichia coli]
1169	Bt1Gc1516	Bt1G2108	2979-3656	g4324613	420	477	2.20E-45	42	100	(AF106566) SlsA [Salmonella typhimurium]
1170	Bt1Gc1515	Bt1G2109	571-1470	g1175720	948	948	2.70E-95	59	100	HYPOTHETICAL OXIDOREDUCTASE IN PBPC-LRPC INTERGENIC REGION [Bacillus subtilis]
1170	Bt1Gc1515	Bt1G2110	2061-1504	g1881232	161	229	4.10E-19	28	99	(AB001488) PROBABLE ACETYLTRANSFERASE. [Bacillus subtilis]
1170	Bt1Gc1515	Bt1G2111	3615-2161	g2462090	2469	2472	8.50E-257	96	100	(Y11171) BC542A protein [Bacillus cereus]
1170	Bt1Gc1515	Bt1G2112	4104-5102	g732353	1015	1003	3.90E-101	59	100	HYPOTHETICAL 36.6 KD PROTEIN IN QOXD-VPR INTERGENIC REGION [Bacillus subtilis]
1171	Bt1Gc1517	Bt1G2114	3770-5142	g1665848	264	374	1.80E-34	23	100	(Z82015) yukC [Bacillus subtilis]
1171	Bt1Gc1517	Bt1G2115	5644-7543	g2104368	266	440	1.00E-41	22	48	(Z95389) hypothetical protein Rv3447c [Mycobacterium tuberculosis]
1171	Bt1Gc1517	Bt1G2116	5846-7543	g1665846	1410	1429	2.80E-146	48	91	(Z82015) yukA [Bacillus subtilis]
1172	Bt1Gc1518	Bt1G2117	974-3406	g586900	3535	3291	0	86	100	NEGATIVE REGULATOR OF GENETIC COMPETENCE MECH [Bacillus subtilis]
1172	Bt1Gc1518	Bt1G2118	3505-4036	g2127105	577	597	4.20E-58	63	68	hypothetical protein Y - Bacillus subtilis (fragment)
1173	Bt1Gc1514	Bt1G2119	1-834	g585209	979	855	1.90E-85	67	62	GLYCEROL-3-PHOSPHATE TRANSPORTER (G-3-P [Bacillus subtilis])

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>1</sub> nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1173	Bt1Gc1514	Bt1G2120	2836-1282	g2226168	121	140	1.10E-09	31	53	TRANSPORTER (G-3-P PERMEASE) [Bacillus subtilis]
1173	Bt1Gc1514	Bt1G2121	3492-1794	g2632425	315	260	8.00E-22	38	34	(Y14080) hypothetical protein [Bacillus subtilis]
1173	Bt1Gc1514	Bt1G2122	4073-4738	g3687663	624	628	2.20E-61	55	100	(Z99104) similar to hypothetical proteins [Bacillus subtilis]
1174	Bt1Gc1519	Bt1G2123	129-569	g2633784	426	426	5.50E-40	56	99	(AF049873) response protein [Lactococcus lactis]
1174	Bt1Gc1519	Bt1G2124	780-1643	g2632234	823	858	9.10E-86	55	100	(Z99111) ykuL [Bacillus subtilis]
1174	Bt1Gc1519	Bt1G2125	1709-2416	g2632238	975	699	6.50E-69	80	100	(AJ222587) YkuM protein [Bacillus subtilis]
1174	Bt1Gc1519	Bt1G2126	2495-3616	g2632239	1244	1175	2.30E-119	63	100	(AJ222587) YkuQ protein [Bacillus subtilis]
1174	Bt1Gc1519	Bt1G2127	3863-5794	g4514629	195	151	5.60E-17	28	83	(AB016894) bd-type quinol oxidase subunit II [Bacillus stearothermophilus]
1174	Bt1Gc1519	Bt1G2128	4830-5492	g732401	824	830	8.50E-83	73	100	HYPOTHETICAL 24.3 KD PROTEIN IN KINC-ADEC INTERGENIC REGION (ORF4) [Bacillus subtilis]
1174	Bt1Gc1519	Bt1G2129	7197-5533	g2833392	2561	2491	8.20E-259	87	100	HYPOTHETICAL 61.5 KD PROTEIN IN ADEC-PDHA INTERGENIC REGION [Bacillus subtilis]
1174	Bt1Gc1519	Bt1G2130	7414-7208	g2633825	202	202	3.00E-16	57	99	(Z99111) ykzG [Bacillus subtilis]
1174	Bt1Gc1519	Bt1G2131	7962-8166	g2633826	208	224	1.40E-18	57	26	(Z99111) similar to hypothetical proteins [Bacillus subtilis]
1175	Bt1Gc1511	Bt1G2132	1679-1	g2635222	228	328	2.90E-29	29	58	(Z99118) similar to N-acetylmuramoyl-L-alanine amidase [Bacillus subtilis]
1175	Bt1Gc1511	Bt1G2133	2746-1893	g3121979	559	604	7.50E-59	45	100	D-ALANINE AMINOTRANSFERASE (D-ASPARTATE AMINOTRANSFERASE) (D-AMINO ACID AMINOTRANSFERASE) (D-AMINO ACID TRANSAMINASE) [Bacillus subtilis]
1175	Bt1Gc1511	Bt1G2134	5262-3415	g1731026	1653	1585	8.40E-163	51	95	HYPOTHETICAL 73.2 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
1176	Bt1Gc1524	Bt1G2135	1-271	g2634117	276	306	2.90E-27	63	30	(Z99113) tRNA isopentenylpyrophosphate transferase [Bacillus subtilis]
1176	Bt1Gc1524	Bt1G2136	314-535	g2634118	319	337	1.50E-30	85	99	(Z99113) similar to host factor-1 protein [Bacillus subtilis]
1176	Bt1Gc1524	Bt1G2137	3780-2836	g4894234	897	900	3.20E-90	54	100	(AF065404) pXO1-18 [Bacillus anthracis]
1176	Bt1Gc1524	Bt1G2138	3888-4847	g134777	1119	1141	9.40E-116	68	100	STAGE V SPORULATION PROTEIN K [Bacillus subtilis]
1176	Bt1Gc1524	Bt1G2139	5739-5134	g2635859	674	561	2.70E-54	62	100	(Z99121) similar to hypothetical proteins [Bacillus subtilis]
1176	Bt1Gc1524	Bt1G2140	5907-6239	g1750108	269	251	1.90E-21	55	33	(U66480) YnbA [Bacillus subtilis]
1177	Bt1Gc1520	Bt1G2141	1-347	g118334	359	357	1.10E-32	62	23	LYSINE DECARBOXYLASE (LDC) [Bacillus subtilis]
1177	Bt1Gc1520	Bt1G2142	426-1166	g2851451	397	415	8.00E-39	35	100	TGL PROTEIN [Bacillus subtilis]
1177	Bt1Gc1520	Bt1G2143	4478-3069	g118672	1964	1833	4.40E-189	79	100	LIPAMIDE DEHYDROGENASE COMPONENT OF

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1177	Bt1Gc1520	Bt1G2144	5772-4487	g129054	1564	1486	2.60E-152	75	100	PYRUVATE DEHYDROGENASE COMPLEX (E3) (DIHYDROLIPOAMIDE DEHYDROGENASE) (S COMPLEX, 50 KD SUBUNIT) [Bacillus subtilis]
1177	Bt1Gc1520	Bt1G2145	6843-5869	g129068	1477	1428	3.60E-146	87	100	DIHYDROLIPOAMIDE ACETYLTRANSFERASE COMPONENT OF PYRUVATE DEHYDROGENASE COMPLEX (E2) (S COMPLEX, 48 KD SUBUNIT) [Bacillus subtilis]
1177	Bt1Gc1520	Bt1G2146	7361-6850	g3123238	739	642	7.10E-63	84	46	PYRUVATE DEHYDROGENASE E1 COMPONENT, BETA SUBUNIT (S COMPLEX, 36 KD SUBUNIT) [Bacillus subtilis]
1178	Bt1Gc1521	Bt1G2147	1-241	g1731031	200	223	1.80E-18	56	38	PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT (S COMPLEX, 42 KD SUBUNIT) (VEGETATIVE PROTEIN 220) (VEG220) [Bacillus subtilis]
1178	Bt1Gc1521	Bt1G2148	180-2393	g1001535	240	375	1.40E-34	26	95	HYPOTHETICAL 23.2 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
1178	Bt1Gc1521	Bt1G2149	1794-1495	g2642588	154	181	5.00E-14	37	99	(D64000) hypothetical protein [Synechocystis PCC6803]
1178	Bt1Gc1521	Bt1G2150	4075-2906	g2633542	1248	1263	1.10E-128	61	100	PagR protein [Bacillus anthracis]
1178	Bt1Gc1521	Bt1G2151	5178-4062	g2633541	1248	1269	2.60E-129	65	100	similar to cystathionine beta-lyase [Bacillus subtilis]
1179	Bt1Gc1522	Bt1G2152	200-965	g1731106	727	685	2.00E-67	55	100	(Z99110) similar to cystathionine gamma-synthase [Bacillus subtilis]
1179	Bt1Gc1522	Bt1G2153	1054-1689	g2634957	798	803	6.20E-80	73	100	HYPOTHETICAL 29.3 KD PROTEIN IN BEX-DNAG/DNAE INTERGENIC REGION (ORF3) [Bacillus subtilis]
1179	Bt1Gc1522	Bt1G2154	1716-2524	g1731001	954	899	4.10E-90	71	100	similar to hypothetical proteins [Bacillus subtilis]
1179	Bt1Gc1522	Bt1G2155	2896-3336	g1731102	158	209	5.40E-17	34	99	HYPOTHETICAL 30.3 KD PROTEIN IN GLYS-DNAG/DNAE INTERGENIC REGION [Bacillus subtilis]
1179	Bt1Gc1522	Bt1G2156	3389-5181	g130904	1422	1412	1.80E-144	48	100	HYPOTHETICAL 17.1 KD PROTEIN IN DNAG/DNAE 5'REGION (P17) [Listeria monocytogenes]
1179	Bt1Gc1522	Bt1G2157	5257-6135	g133466	1246	1273	9.60E-130	84	79	DNA PRIMASE [Bacillus subtilis]
1180	Bt1Gc1526	Bt1G2158	1455-2276	g1620924	530	320	9.40E-29	38	100	RNA POLYMERASE SIGMA FACTOR RPOD (SIGMA-A) (SIGMA-43) [Bacillus subtilis]
1180	Bt1Gc1526	Bt1G2159	1227-4443	g2633724	579	640	1.20E-62	38	52	(Z79580) putative ORF [Bacillus subtilis]
										similar to two-component sensor histidine kinase



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1180	Bt1Gc1526	Bt1G2160	5089-4484	g2415393	654	668	1.20E-65	60	100	[Bacillus subtilis] YodD [AF015775]
1180	Bt1Gc1526	Bt1G2161	5965-5120	g1881356	1280	1280	1.70E-130	82	90	(AB001488) SIMILAR TO PENTACHLOROPHENOL-INDUCED PERIPLASMIC PROTEIN IN FLAVOBACTERIUM SP [Bacillus subtilis] (AB001488) SIMILAR TO PYRUVATE OXIDASE AND ACETOLACTATE SYNTHASE. [Bacillus subtilis] (AE000757) hypothetical protein [Aquifex aeolicus] (AC006921) hypothetical protein [Arabidopsis thaliana] (AJ006589) gp35 [Bacteriophage phi-C31] (AJ000394) transcriptional regulatory protein CelR [Bacillus cereus] (AJ000394) hypothetical protein [Bacillus cereus] (AJ000394) chromate transport protein [Bacillus cereus] (AJ000394) yjC-like protein [Bacillus cereus] PROBABLE 6-PHOSPHO-BETA-GLUCOSIDASE [Bacillus subtilis] (D87979) YfmJ [Bacillus subtilis] ALCOHOL DEHYDROGENASE I (ADH I) [Zymomonas mobilis] (HYPOTHETICAL 32.3 KD PROTEIN IN RHSE-NARV INTERGENIC REGION (ORFB) [Escherichia coli] HYPOTHETICAL 31.2 KD PROTEIN IN NAGH 5'REGION (ORFB) [Clostridium perfringens] (Y14079) hypothetical protein [Bacillus subtilis] (Y14079) hypothetical protein [Bacillus subtilis] (D83967) YfkC [Bacillus subtilis] 18 KD HEAT SHOCK PROTEIN (HSP 18) [Clostridium acetobutylicum] (Z99120) similar to pyrazinamidase/nicotinamidase [Bacillus subtilis] (U17283) putative spore germination apparatus protein [Bacillus megaterium]
1181	Bt1Gc1525	Bt1G2162	4205-2478	g1881244	1621	1551	3.30E-159	54	100	
1181	Bt1Gc1525	Bt1G2163	5069-4358	g2984109	89	146	3.50E-10	23	100	
1181	Bt1Gc1525	Bt1G2164	7779-7200	g4510346	60	153	1.40E-10	25	61	
1182	Bt1Gc1527	Bt1G2165	4351-3712	g3947454	219	276	4.30E-24	32	100	
1183	Bt1Gc1530	Bt1G2166	1-200	g4584200	337	354	2.30E-32	99	14	
1183	Bt1Gc1530	Bt1G2167	921-511	g4584201	598	455	4.60E-43	83	99	
1183	Bt1Gc1530	Bt1G2168	1109-2287	g4584202	1903	1674	3.10E-172	93	100	
1183	Bt1Gc1530	Bt1G2169	3033-2332	g4584203	1150	1036	1.30E-104	94	100	
1183	Bt1Gc1530	Bt1G2170	4267-3086	g1168884	1603	1604	8.10E-165	77	89	
1184	Bt1Gc1529	Bt1G2171	1003-2275	g3915990	1152	1084	1.00E-109	52	79	
1184	Bt1Gc1529	Bt1G2172	3072-6263	g2116974	4203	4166	0	76	100	
1185	Bt1Gc1528	Bt1G2173	861-1	g113368	1034	1064	1.40E-107	70	84	
1185	Bt1Gc1528	Bt1G2174	2012-574	g2506666	212	301	9.70E-27	29	100	
1185	Bt1Gc1528	Bt1G2175	3466-2186	g141086	340	366	1.20E-33	32	100	
1185	Bt1Gc1528	Bt1G2176	3871-3230	g2226141	509	496	2.10E-47	49	100	
1185	Bt1Gc1528	Bt1G2177	5004-3871	g2226140	888	839	9.40E-84	49	100	
1185	Bt1Gc1528	Bt1G2178	6049-5210	g2626815	538	427	4.30E-40	42	100	
1185	Bt1Gc1528	Bt1G2179	6804-6365	g462323	210	234	1.20E-19	37	99	
1185	Bt1Gc1528	Bt1G2180	7699-7151	g2635671	670	675	2.30E-66	66	99	
1186	Bt1Gc1531	Bt1G2181	190-1344	g1098508	1190	982	6.60E-99	58	100	

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1186	Bt1Gc1531	Bt1G2182	3141-1777	g2828205	1579	1552	2.60E-159	66	100	PUTATIVE AMINOTRANSFERASE YHXA [Bacillus subtilis]
1186	Bt1Gc1531	Bt1G2183	4875-3327	g629038	1866	1504	3.20E-154	74	100	hypothetical protein - Bacillus subtilis []
1186	Bt1Gc1531	Bt1G2184	5176-4871	g732347	243	119	1.90E-07	41	99	HYPOTHETICAL 11.7 KD PROTEIN IN EPR-GALK INTERGENIC REGION [Bacillus subtilis]
1186	Bt1Gc1531	Bt1G2185	7322-4429	g3219336	81	224	1.40E-18	30	50	(AC004685) Unknown gene product [Homo sapiens]
1186	Bt1Gc1531	Bt1G2186	7082-6453	g2632595	420	447	3.30E-42	40	100	(Z99105) ycgF [Bacillus subtilis]
1186	Bt1Gc1531	Bt1G2187	7940-7199	g1805382	161	254	9.20E-22	26	100	(D50453) ycgG [Bacillus subtilis]
1187	Bt1Gc1534	Bt1G2188	141-662	g1651966	167	190	5.60E-15	29	99	(D90901) hypothetical protein [Synechocystis sp.]
1187	Bt1Gc1534	Bt1G2189	1957-3118	g1002992	316	416	6.30E-39	32	100	(U34772) ORF375 [Dichelobacter nodosus]
1187	Bt1Gc1534	Bt1G2190	2698-3087	g1149666	231	229	4.10E-19	47	99	(X86498) IS1136 DNA [Clostridium perfringens]
1187	Bt1Gc1534	Bt1G2191	3239-4822	g2495662	1847	1443	9.30E-148	67	95	PUTATIVE L-LACTATE PERMEASE [Escherichia coli]
1188	Bt1Gc1532	Bt1G2192	554-264	g134776	454	420	2.40E-39	91	99	STAGE V SPOULATION PROTEIN G [Bacillus subtilis]
1188	Bt1Gc1532	Bt1G2193	1084-710	g586881	467	469	1.50E-44	74	99	HYPOTHETICAL 13.7 KD PROTEIN IN PURR-SPOVG INTERGENIC REGION (ORF2) [Bacillus subtilis]
1188	Bt1Gc1532	Bt1G2194	2049-1195	g586880	933	945	5.50E-95	65	100	PUR OPERON REPRESSOR [Bacillus subtilis]
1188	Bt1Gc1532	Bt1G2195	2973-2107	g586879	1130	1001	6.40E-101	74	100	HYPOTHETICAL 31.7 KD PROTEIN IN SSPF-PURR INTERGENIC REGION (ORF1) [Bacillus subtilis]
1188	Bt1Gc1532	Bt1G2196	3349-3173	g1711546	296	296	3.30E-26	98	98	SSPF PROTEIN [Bacillus cereus]
1188	Bt1Gc1532	Bt1G2197	3701-3444	g586231	283	284	6.10E-25	65	99	VEG PROTEIN [Bacillus subtilis]
1188	Bt1Gc1532	Bt1G2198	4790-3936	g586878	902	838	1.20E-83	59	100	HYPOTHETICAL 33.3 KD PROTEIN IN KSGA-VEG INTERGENIC REGION [Bacillus subtilis]
1188	Bt1Gc1532	Bt1G2199	5782-4907	g585375	1044	1054	1.60E-106	70	100	DIMETHYLADENOSINE TRANSFERASE (S-ADENOSYLMETHIONINE-6-N', N'-ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RRNA)
1188	Bt1Gc1532	Bt1G2200	6333-5776	g586877	659	663	4.20E-65	69	99	DIMETHYLTRANSFERASE (HIGH LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN DIMETHYLTRANSFERASE) [Bacillus subtilis]
1189	Bt1Gc1533	Bt1G2201	165-1	g134763	167	119	8.20E-07	58	17	HYPOTHETICAL 20.7 KD PROTEIN IN METS-KSGA INTERGENIC REGION [Bacillus subtilis]
1189	Bt1Gc1533	Bt1G2202	1571-417	g120577	1421	1162	3.70E-122	78	100	SPOULATION SIGMA-E FACTOR PROCESSING PEPTIDASE (STAGE II SPOULATION PROTEIN GA) [Bacillus subtilis]
1189	Bt1Gc1533	Bt1G2203	2918-1635	g120567	1265	1223	1.90E-124	57	100	CELL DIVISION PROTEIN FTSZ [Bacillus subtilis]
1189	Bt1Gc1533	Bt1G2203	2918-1635	g120567	1265	1223	1.90E-124	57	100	CELL DIVISION PROTEIN FTSA [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat-nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1189	Bt1Gc1533	Bt1G2204	4100-3303	g118667	389	378	6.70E-35	35	100	DIVISION INITIATION PROTEIN (CELL DIVISION AND SPORULATION PROTEIN) [Bacillus subtilis]
1189	Bt1Gc1533	Bt1G2205	5092-4184	g140583	1086	1089	3.00E-110	66	100	UDP-N-ACETYLENOLPYRUVOYLGLUCOSAMINE REDUCTASE (UDP-N-ACETYLMURAMATE DEHYDROGENASE) [Bacillus subtilis]
1189	Bt1Gc1533	Bt1G2206	6399-5311	g585531	1190	1146	2.80E-116	62	100	UDP-N-ACETYLMURAMATE--N-ACETYLMURAMYL-(PENTAPEPTIDE) PYROPHOSPHORYL-UNDECAPRENOL N-ACETYLMURAMINE TRANSFERASE [Bacillus subtilis]
1189	Bt1Gc1533	Bt1G2207	7375-6506	g580938	1251	865	1.70E-86	81	100	(X51419) internal open reading frame (AA 1-290) [Bacillus subtilis]
1190	Bt1Gc1535	Bt1G2208	18-985	g2337837	434	461	1.10E-43	37	100	(Z98604) hypothetical protein MLCB2052.31 [Mycobacterium leprae]
1190	Bt1Gc1535	Bt1G2209	2742-1296	g1881344	1233	1258	3.70E-128	50	100	(AB001488) SIMILAR TO THE RHIZOPINE CATABOLISM (MOCR) GENE OF RHIZOBIUM MELILOTI. [Bacillus subtilis]
1190	Bt1Gc1535	Bt1G2210	2821-3479	g2145387	479	490	9.10E-47	47	100	(Y09476) YisU [Bacillus subtilis]
1190	Bt1Gc1535	Bt1G2211	4774-3504	g2633434	931	787	3.10E-78	42	100	(Z99109) similar to multidrug resistance protein [Bacillus subtilis]
1190	Bt1Gc1535	Bt1G2212	5213-4995	g134241	259	267	3.90E-23	71	99	SMALL, ACID-SOLUBLE SPORE PROTEIN C5 (SASP) [Bacillus megaterium]
1190	Bt1Gc1535	Bt1G2213	10092-7477	g2648777	106	174	2.80E-12	28	28	(AE000980) long-chain-fatty-acid--CoA ligase (fadD-7) [Archaeoglobus fulgidus]
1190	Bt1Gc1535	Bt1G2214	9948-8717	g4151936	370	496	2.10E-47	29	68	(AF110737) RhsF [Sinorhizobium meliloti]
1190	Bt1Gc1535	Bt1G2215	10032-10265	g2497382	286	292	9.10E-26	76	18	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1191	Bt1Gc1538	Bt1G2216	1-565	g2634202	616	623	7.30E-61	59	95	(Z99113) similar to alkaline phosphatase [Bacillus subtilis]
1191	Bt1Gc1538	Bt1G2217	1226-1678	g2632052	101	116	3.90E-07	25	99	(AJ002571) YkoM [Bacillus subtilis]
1191	Bt1Gc1538	Bt1G2218	1706-2401	g1176281	387	416	6.30E-39	39	100	HYPOTHETICAL 26.3 KD PROTEIN IN GNTR-GGT INTERGENIC REGION (F231) [Escherichia coli]
1191	Bt1Gc1538	Bt1G2219	2564-4207	g730171	1386	1471	1.00E-150	51	100	NEUTRAL PROTEASE B PRECURSOR [Bacillus subtilis]
1191	Bt1Gc1538	Bt1G2220	4602-4294	g1620928	301	224	1.40E-18	58	99	(Z79580) putative orf [Bacillus subtilis]
1192	Bt1Gc1536	Bt1G2221	523-1	g2619033	673	634	5.00E-62	77	20	(AF027868) PEP synthase [Bacillus subtilis]
1192	Bt1Gc1536	Bt1G2222	1932-1309	g2632595	351	365	1.60E-33	33	100	(Z99105) ycgF [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1192	Bt1Gc1536	Bt1G2223	2790-3215	g2632725	212	241	2.20E-20	33	99	(Z99106) transcriptional regulator (Lrp/AsnC family) [Bacillus subtilis]
1192	Bt1Gc1536	Bt1G2224	3343-4084	g2492749	466	536	1.20E-51	45	100	3-OXOACYL- [Thermotoga maritima]
1192	Bt1Gc1536	Bt1G2225	4897-6112	g2394395	662	641	9.00E-63	51	100	(AF017435) orfL1 [Methylobacterium extorquens]
1192	Bt1Gc1536	Bt1G2226	6970-6458	g1945676	526	433	9.90E-41	56	99	(Z94043) hypothetical protein [Bacillus subtilis]
1192	Bt1Gc1536	Bt1G2227	7478-7256	g134229	214	249	3.10E-21	75	99	SMALL, ACID-SOLUBLE SPORE PROTEIN 1 (SASP) [Sporosarcina halophila]
1192	Bt1Gc1536	Bt1G2228	7744-8874	g2108273	1284	1246	7.00E-127	62	100	(X92868) NAD alcohol dehydrogenase [Bacillus subtilis]
1192	Bt1Gc1536	Bt1G2229	10228-9146	g2633724	328	351	3.80E-31	50	20	(Z99111) similar to two-component sensor histidine kinase [Bacillus subtilis]
1193	Bt1Gc1539	Bt1G2230	1-1505	g2495662	1226	987	2.00E-99	50	92	PUTATIVE L-LACTATE PERMEASE [Escherichia coli]
1193	Bt1Gc1539	Bt1G2231	2969-1545	g114271	1902	1801	1.10E-185	77	100	ASPARTATE AMMONIA-LYASE (ASPARTASE) [Bacillus subtilis]
1193	Bt1Gc1539	Bt1G2232	4438-3110	g3025180	1303	1078	4.40E-109	57	100	HYPOTHETICAL 49.9 KD PROTEIN IN CITA-SSPB INTERGENIC REGION [Bacillus subtilis]
1193	Bt1Gc1539	Bt1G2233	4933-4550	g1176959	420	422	1.50E-39	62	99	HYPOTHETICAL 14.8 KD PROTEIN IN TDK-PRFA INTERGENIC REGION [Bacillus subtilis]
1193	Bt1Gc1539	Bt1G2234	5453-5156	g225559	478	494	3.40E-47	99	21	ORF IS231C [Bacillus thuringiensis]
1194	Bt1Gc1540	Bt1G2235	937-1257	g2633561	178	167	1.50E-12	38	99	(Z99110) yjdJ [Bacillus subtilis]
1194	Bt1Gc1540	Bt1G2236	2579-1320	g135192	1591	1617	3.40E-166	71	100	TYROSYL-TRNA SYNTHETASE 1 (TYROSINE--TRNA LIGASE) (TYRRS 1) [Bacillus subtilis]
1194	Bt1Gc1540	Bt1G2237	3717-3202	g1256141	125	194	2.10E-15	26	99	(AB002150) YbbL [Bacillus subtilis]
1194	Bt1Gc1540	Bt1G2238	4786-4232	g586845	650	581	2.10E-56	66	99	PROBABLE MALTOSE O-ACETYLTRANSFERASE (MALTOSE TRANSACETYLASE) [Bacillus subtilis]
1194	Bt1Gc1540	Bt1G2239	5962-4874	g728788	1623	1659	1.20E-170	82	63	ACETYL-COENZYME A SYNTHETASE (ACETATE--COA LIGASE) (ACYL-ACTIVATING ENZYME) (ACETYL-COA SYNTHASE) [Bacillus subtilis]
1195	Bt1Gc1537	Bt1G2240	479-2128	g1075696	802	879	5.40E-88	36	100	probable pheromone binding protein prgZ - Enterococcus faecalis plasmid pCF10 [Plasmid pCF10]
1195	Bt1Gc1537	Bt1G2241	2263-3195	g2633498	805	658	1.40E-64	50	100	(Z99110) oligopeptide ABC transporter (permease) [Bacillus subtilis]
1195	Bt1Gc1537	Bt1G2242	3195-4175	g1732315	823	628	2.20E-61	51	100	(U78885) transport system permease homolog [Listeria monocytogenes]
1195	Bt1Gc1537	Bt1G2243	4215-4794	g2633500	684	688	9.40E-68	70	54	(Z99110) oligopeptide ABC transporter (ATP-binding protein) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat- nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1196	Bt1Gc1541	Bt1G2244	1277-102	g2415747	969	844	2.80E-84	49	100	(AB000617) YceI [Bacillus subtilis]
1196	Bt1Gc1541	Bt1G2245	2260-1331	g1176985	719	734	1.30E-72	45	100	IOLS PROTEIN (VEGETATIVE PROTEIN 147) (VEG147) [Bacillus subtilis]
1196	Bt1Gc1541	Bt1G2246	2367-2675	g2632689	245	269	2.40E-23	49	99	(Z99106) similar to transcriptional regulator (ArsR family) [Bacillus subtilis]
1196	Bt1Gc1541	Bt1G2247	3643-2675	g2118109	974	989	1.20E-99	58	100	microbial serine proteinase (EC 3.4.21.-) ispQ - Bacillus sp [Bacillus sp.]
1196	Bt1Gc1541	Bt1G2248	4226-3945	g39805	258	259	2.70E-22	53	99	(X61953) abrB [Bacillus subtilis]
1196	Bt1Gc1541	Bt1G2249	6222-4360	g1731300	910	762	1.40E-75	36	100	HYPOTHETICAL 70.5 KD PROTEIN IN IDH 3'REGION [Bacillus subtilis]
1196	Bt1Gc1541	Bt1G2250	6498-7313	g129288	1099	1099	2.60E-111	78	100	NH(3)-DEPENDENT NAD(+) SYNTHETASE (SPORE OUTGROWTH FACTOR B) (SPORULATION PROTEIN OUTB) (GENERAL STRESS PROTEIN 38) (GSP38) [Bacillus subtilis]
1196	Bt1Gc1541	Bt1G2251	7761-7369	g1177029	173	233	4.50E-19	36	33	HYPOTHETICAL 45.4 KD PROTEIN IN THIAMINASE I 5'REGION []
1197	Bt1Gc1542	Bt1G2252	1-785	g3913073	386	434	7.80E-41	38	88	PROBABLE 2-DEHYDRO-PANTOATE 2-REDUCTASE (KETOPANTOATE REDUCTASE) (KPA REDUCTASE) [Bacillus subtilis]
1197	Bt1Gc1542	Bt1G2253	856-2469	g2633883	1085	1123	7.60E-114	41	100	(Z99111) yIIA [Bacillus subtilis]
1197	Bt1Gc1542	Bt1G2254	2842-3774	g1730600	1268	1202	3.20E-122	79	100	HYPOTHETICAL 35.3 KD PROTEIN IN FTSL 5'REGION (ORFB) [Bacillus subtilis]
1197	Bt1Gc1542	Bt1G2255	3771-4170	g2149892	149	176	1.70E-13	27	99	(U94706) cell division protein [Staphylococcus aureus]
1197	Bt1Gc1542	Bt1G2256	4171-6294	g585648	1623	1691	4.90E-174	47	100	PENICILLIN-BINDING PROTEIN 2B (PBP-2B) [Bacillus subtilis]
1197	Bt1Gc1542	Bt1G2257	6406-8343	g586022	2382	2318	1.80E-240	69	100	STAGE V SPORULATION PROTEIN D (SPORULATION SPECIFIC PENICILLIN-BINDING PROTEIN) [Bacillus subtilis]
1198	Bt1Gc1544	Bt1G2258	2485-101	g730776	2423	1822	1.90E-225	63	100	STAGE III SPORULATION PROTEIN E [Bacillus subtilis]
1198	Bt1Gc1544	Bt1G2259	3635-2901	g143581	930	949	2.10E-95	75	100	(M17445) ORF X [Bacillus subtilis]
1198	Bt1Gc1544	Bt1G2260	5300-3753	g2634050	1652	1560	3.70E-160	60	100	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
1198	Bt1Gc1544	Bt1G2261	7032-6166	g416876	926	814	4.20E-81	61	100	DIHYDRODIPICOLINATE SYNTHASE (DHDPS) (VEGETATIVE PROTEIN 81) (VEG81) [Bacillus subtilis]
1198	Bt1Gc1544	Bt1G2262	8279-7068	g416595	1382	1382	2.70E-141	67	100	ASPARTOKINASE 1 ALPHA AND BETA SUBUNITS (ASPARTATE KINASE 1) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1199	Bt1Gc1545	Bt1G2263	377-1	g1731101	462	474	4.50E-45	74	46	HYPOTHETICAL 29.7 KD PROTEIN IN FOLD-AHRC INTERGENIC REGION [Bacillus subtilis]
1199	Bt1Gc1545	Bt1G2264	2273-376	g1731052	2603	2568	5.70E-267	77	100	PROBABLE 1-DEOXYXYLULOSE-5-PHOSPHATE SYNTHASE (DXP SYNTHASE) [Bacillus subtilis]
1199	Bt1Gc1545	Bt1G2265	3468-2578	g585326	865	809	1.40E-80	57	100	GERANYL TRANSFERASE (FARNESYL-DIPHOSPHATE SYNTHASE) (FPP SYNTHASE) [Bacillus stearothermophilus]
1199	Bt1Gc1545	Bt1G2266	3620-3459	g1706724	147	154	3.60E-11	54	98	PUTATIVE EXODEOXYRIBONUCLEASE SMALL SUBUNIT (EXONUCLEASE VII SMALL SUBUNIT) [Bacillus subtilis]
1199	Bt1Gc1545	Bt1G2267	5046-3700	g1706723	1226	1168	1.30E-118	54	100	PUTATIVE EXODEOXYRIBONUCLEASE LARGE SUBUNIT (EXONUCLEASE VII LARGE SUBUNIT) [Bacillus subtilis]
1199	Bt1Gc1545	Bt1G2268	5930-5082	g1706886	1003	975	3.70E-98	67	100	METHYLENETETRAHYDROFOLATE DEHYDROGENASE / METHENYL TETRAHYDROFOLATE CYCLOHYDROLASE [Bacillus subtilis]
1199	Bt1Gc1545	Bt1G2269	6346-5948	g1709418	342	378	6.70E-35	54	99	N UTILIZATION SUBSTANCE PROTEIN B HOMOLOG (NUSB PROTEIN) [Bacillus subtilis]
1200	Bt1Gc1543	Bt1G2270	1-8387	g2982196	3841	4102	0	35	43	(AF007865) bacitracin synthetase 3; BacC [Bacillus licheniformis]
1201	Bt1Gc1546	Bt1G2271	1-354	g4160472	399	383	2.00E-35	64	80	(AF109909) PhaQ [Bacillus megaterium]
1201	Bt1Gc1546	Bt1G2272	1071-1472	g2335053	267	273	9.00E-24	41	99	(D88825) ORF3 [Aeromonas caviae]
1201	Bt1Gc1546	Bt1G2273	1626-1820	g134237	229	238	4.60E-20	72	98	SMALL, ACID-SOLUBLE SPORE PROTEIN C3 (SASP) [Bacillus megaterium]
1201	Bt1Gc1546	Bt1G2274	1911-3242	g2635594	1192	916	6.50E-92	53	100	(Z99119) similar to Na <sup>+</sup> -transporting ATP synthase [Bacillus subtilis]
1201	Bt1Gc1546	Bt1G2275	3541-5356	g4160468	956	804	4.80E-80	73	96	(AF109909) putative protein [Bacillus megaterium]
1201	Bt1Gc1546	Bt1G2276	4538-5305	g732355	641	528	8.50E-51	48	100	HYPOTHETICAL 28.4 KD PROTEIN IN SACT-SACP INTERGENIC REGION [Bacillus subtilis]
1201	Bt1Gc1546	Bt1G2277	6460-5717	g116906	356	311	8.40E-28	35	100	TYPE 4 PREPILIN-LIKE PROTEIN SPECIFIC LEADER PEPTIDASE (LATE COMPETENCE PROTEIN COMC) [Bacillus subtilis]
1201	Bt1Gc1546	Bt1G2278	7292-6573	g2635916	603	609	2.20E-59	48	100	(Z99121) similar to hypothetical proteins [Bacillus subtilis]
1201	Bt1Gc1546	Bt1G2279	8710-7280	g2635917	1809	1762	1.50E-181	71	100	(Z99121) similar to hypothetical proteins [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvr	NCBI gi description
1201	Bt1Gc1546	Bt1G2280	9446-8733	g1945722	950	950	1.60E-95	74	100	(Z94043) hypothetical protein [Bacillus subtilis]
1201	Bt1Gc1546	Bt1G2281	9973-9615	g2635931	230	215	1.30E-17	43	68	(Z99121) similar to transcriptional regulator (GntR family) [Bacillus subtilis]
1202	Bt1Gc1547	Bt1G2282	3544-1338	g2493678	97	218	7.70E-17	25	35	OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN) (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN) [Mus musculus]
1202	Bt1Gc1547	Bt1G2283	4729-1097	g2444121	83	192	7.70E-14	36	10	(U88974) ORF42 [Streptococcus thermophilus temperate bacteriophage O1205]
1203	Bt1Gc1523	Bt1G2284	7-1437	g2494077	1338	1354	2.50E-138	55	100	NADP-DEPENDENT GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (NON-PHOSPHORYLATING GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE) (GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (NADP+)) (TRIOSEPHOSPHATE DEHYDROGENASE) [Streptococcus mutans]
1203	Bt1Gc1523	Bt1G2285	4524-2737	g1709184	91	186	1.50E-14	37	40	GLUTAMATE RACEMASE [Bacillus sphaericus]
1203	Bt1Gc1523	Bt1G2286	3391-4797	g2635847	1477	1483	5.40E-152	59	100	(Z99121) similar to ABC transporter (amino acid permease) [Bacillus subtilis]
1203	Bt1Gc1523	Bt1G2287	4876-6074	g1706578	469	448	2.60E-42	29	100	PROTEIN ECSB [Bacillus subtilis]
1203	Bt1Gc1523	Bt1G2288	6248-6453	g231698	153	195	9.10E-15	59	15	CATALASE []
1204	Bt1Gc1550	Bt1G2289	327-1	g225559	503	515	2.00E-49	94	23	ORF IS231C [Bacillus thuringiensis]
1204	Bt1Gc1550	Bt1G2290	717-4011	g128494	67	148	3.30E-09	27	12	NODULATION PROTEIN V [Bradyrhizobium japonicum]
1204	Bt1Gc1550	Bt1G2291	1-2399	g2633696	816	934	8.10E-94	35	76	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
1204	Bt1Gc1550	Bt1G2292	1062-3089	g1763702	53	147	1.00E-08	24	69	(Z83337) similar to phosphatases [Bacillus subtilis]
1205	Bt1Gc1549	Bt1G2293	2777-3964	g585035	1214	1145	3.50E-116	62	100	PENICILLIN-BINDING PROTEIN DACF PRECURSOR (D-ALANYL-D-ALANINE CARBOXYPEPTIDASE) (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) [Bacillus subtilis]
1205	Bt1Gc1549	Bt1G2294	4137-4484	g3287912	431	431	1.60E-40	72	99	ANTI-SIGMA F FACTOR ANTAGONIST (STAGE II SPORULATION PROTEIN AA) [Bacillus coagulans]
1205	Bt1Gc1549	Bt1G2295	4488-4925	g134757	591	591	1.80E-57	79	99	ANTI-SIGMA F FACTOR (STAGE II SPORULATION PROTEIN AB) [Bacillus licheniformis]
1205	Bt1Gc1549	Bt1G2296	4941-5699	g464690	1071	1023	3.00E-103	83	100	RNA POLYMERASE SIGMA-F FACTOR (STAGE II SPORULATION PROTEIN AC) (SPORULATION SIGMA

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1205	Bt1Gc1549	Bt1G22297	5832-7304	g2648543	1132	1006	1.90E-101	47	100	FACTOR) [Bacillus megaterium] (AE000965) sodium- and chloride-dependent transporter [Archaeoglobus fulgidus]
1205	Bt1Gc1549	Bt1G22298	7572-8172	g730782	419	298	2.00E-26	43	100	STAGE V SPOULATION PROTEIN AA [Bacillus subtilis]
1206	Bt1Gc1548	Bt1G22299	1965-554	g728922	2035	1969	1.70E-203	86	100	ATP SYNTHASE BETA CHAIN [Bacillus caldotenax]
1206	Bt1Gc1548	Bt1G22300	3067-2210	g114636	1052	1070	3.10E-108	72	100	ATP SYNTHASE GAMMA CHAIN [Bacillus megaterium]
1206	Bt1Gc1548	Bt1G22301	4331-3361	g114510	1429	1413	1.40E-144	86	64	ATP SYNTHASE ALPHA CHAIN [Bacillus megaterium]
1207	Bt1Gc1551	Bt1G22302	320-1	g2497382	532	548	6.50E-53	99	25	TRANSPORASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1207	Bt1Gc1551	Bt1G22303	1586-1992	g3257497	63	126	3.40E-08	27	99	(AP000004) 132aa long hypothetical protein [Pyrococcus horikoshii]
1207	Bt1Gc1551	Bt1G22304	4121-4587	g2370586	85	133	6.10E-09	44	25	(Y11313) hypothetical protein [Xanthomonas campestris]
1208	Bt1Gc1552	Bt1G22305	400-1	g2612891	306	311	8.40E-28	41	78	(AF015825) MutT homolog [Bacillus subtilis]
1208	Bt1Gc1552	Bt1G22306	1702-617	g1929335	1142	1149	1.30E-116	56	100	(Z93767) ywrH [Bacillus subtilis]
1208	Bt1Gc1552	Bt1G22307	3069-2433	g1001645	201	213	2.00E-17	28	100	(D64002) hypothetical protein [Synecocystis sp.]
1208	Bt1Gc1552	Bt1G22308	4061-3630	g2634168	444	462	8.40E-44	59	99	(Z99113) similar to fosfomycin resistance protein [Bacillus subtilis]
1208	Bt1Gc1552	Bt1G22309	5782-4124	g468461	761	908	4.60E-91	36	100	(D28859) TraC [Enterococcus faecalis]
1208	Bt1Gc1552	Bt1G22310	8931-6873	g2935421	294	355	1.80E-32	49	38	(AF047839) adaptive response regulatory protein [Pseudoalteromonas sp. S9]
1208	Bt1Gc1552	Bt1G22311	7462-8475	g1731075	1158	1162	5.60E-118	65	100	PROBABLE NADH-DEPENDENT FLAVIN OXIDOREDUCTASE YQJM [Bacillus subtilis]
1208	Bt1Gc1552	Bt1G22312	9233-7520	g3043927	212	259	2.70E-22	34	85	(AF039103) Tat-interacting protein TIP30 [Homo sapiens]
1208	Bt1Gc1552	Bt1G22313	9596-9487	g225559	155	161	5.10E-11	86	8	ORF IS231C [Bacillus thuringiensis]
1209	Bt1Gc1556	Bt1G22314	639-3659	g2622469	1087	1282	1.10E-130	43	97	(AE000899) ferrous iron transport protein B [Methanobacterium thermoautotrophicum]
1209	Bt1Gc1556	Bt1G22315	3833-5640	g3123076	297	276	4.30E-24	34	87	HYPOTHETICAL 28.9 KD PROTEIN SLL0617 [Synecocystis sp.]
1210	Bt1Gc1555	Bt1G22316	218-1021	g2635730	1155	1162	5.60E-118	82	100	(Z99120) similar to lipoic acid synthetase [Bacillus subtilis]
1210	Bt1Gc1555	Bt1G22317	3526-2762	g2226236	422	416	6.30E-39	38	100	(Y14083) hypothetical protein [Bacillus subtilis]
1210	Bt1Gc1555	Bt1G22318	3669-4409	g1592700	478	528	8.50E-51	40	100	(Y08559) Unknown [Bacillus subtilis]
1210	Bt1Gc1555	Bt1G22319	5016-4415	g2635729	189	242	1.70E-20	27	100	(Z99120) yutC [Bacillus subtilis]
1210	Bt1Gc1555	Bt1G22320	5129-5434	g2635728	377	390	3.60E-36	62	99	(Z99120) yutD [Bacillus subtilis]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1210	Bt1Ge1555	Bt1G2321	6176-4268	g2501678	71	130	4.80E-08	27	30	HYPOTHETICAL 33.7 KD PROTEIN IN LSP-PYRR INTERGENIC REGION (ORF-X) []
1211	Bt1Ge1554	Bt1G2322	230-1	g1176915	168	200	4.90E-16	44	52	HYPOTHETICAL 17.3 KD PROTEIN IN LYSC 3 REGION [Bacillus subtilis]
1211	Bt1Ge1554	Bt1G2323	525-1130	g118620	811	819	1.20E-81	71	100	SUCCINATE DEHYDROGENASE CYTOCHROME B-558 SUBUNIT [Bacillus subtilis]
1211	Bt1Ge1554	Bt1G2324	1165-2922	g1071812	2563	2570	3.50E-267	82	100	succinate dehydrogenase (EC 1.3.99.1) flavoprotein - Bacillus subtilis []
1211	Bt1Ge1554	Bt1G2325	2951-3709	g118613	1160	1169	1.00E-118	82	100	SUCCINATE DEHYDROGENASE IRON-SULFUR PROTEIN [Bacillus subtilis]
1211	Bt1Ge1554	Bt1G2326	3808-4258	g140720	353	393	1.70E-36	50	99	HYPOTHETICAL 17.1 KD PROTEIN IN SDHB-GERE INTERGENIC REGION [Bacillus subtilis]
1211	Bt1Ge1554	Bt1G2327	5800-4288	g2529473	508	570	3.00E-55	44	100	(AF006665) YokZ [Bacillus subtilis]
1211	Bt1Ge1554	Bt1G2328	5337-6148	g127247	692	668	1.20E-65	51	100	CHEMOTAXIS MOTA PROTEIN (MOTILITY PROTEIN A) [Bacillus subtilis]
1211	Bt1Ge1554	Bt1G2329	6133-6780	g127250	526	445	5.30E-42	45	82	CHEMOTAXIS MOTB PROTEIN (MOTILITY PROTEIN B) [Bacillus subtilis]
1212	Bt1Ge1553	Bt1G2330	1593-757	g586827	819	775	5.70E-77	55	100	HYPOTHETICAL 31.3 KD PROTEIN IN COTF-TETB INTERGENIC REGION [Bacillus subtilis]
1212	Bt1Ge1553	Bt1G2331	1771-2016	g1176942	137	154	3.60E-11	39	99	HYPOTHETICAL 8.9 KD PROTEIN IN UREA-UREB INTERGENIC REGION (ORF3) [Rhizobium meliloti]
1212	Bt1Ge1553	Bt1G2332	3064-2011	g2632022	1106	1142	7.30E-116	62	100	(AJ002571) YkgB [Bacillus subtilis]
1212	Bt1Ge1553	Bt1G2333	4735-3194	g121516	1712	1706	1.30E-175	64	100	GLUCONOKINASE (GLUCONATE KINASE) [Bacillus subtilis]
1212	Bt1Ge1553	Bt1G2334	6213-4875	g4103625	1017	754	9.60E-75	46	100	(AF026470) gluconate permease [Pseudomonas aeruginosa]
1212	Bt1Ge1553	Bt1G2335	7256-6585	g2982855	511	547	8.30E-53	51	100	(AE000674) transaldolase [Aquifex aeolicus]
1212	Bt1Ge1553	Bt1G2336	7963-7369	g1730976	842	844	2.80E-84	82	67	HYPOTHETICAL 32.8 KD PROTEIN IN NUCB-AROD INTERGENIC REGION [Bacillus subtilis]
1213	Bt1Ge1557	Bt1G2337	834-391	g586818	291	296	3.30E-26	37	99	HYPOTHETICAL 16.0 KD PROTEIN IN COTF-TETB INTERGENIC REGION [Bacillus subtilis]
1213	Bt1Ge1557	Bt1G2338	1903-1346	g280166	407	421	1.90E-39	44	99	rfo protein - Vibrio cholerae [Vibrio cholerae]
1213	Bt1Ge1557	Bt1G2339	3122-1896	g2496480	397	502	4.80E-48	30	100	HYPOTHETICAL 44.9 KD PROTEIN CY78.26 [Mycobacterium tuberculosis]
1213	Bt1Ge1557	Bt1G2340	8715-6985	g2829799	1397	1379	5.60E-141	47	100	TRANSPORT ATP-BINDING PROTEIN CYDD [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1213	Bt1Gc1557	Bt1G2341	10060-8706	g2829798	1184	1104	7.80E-112	57	79	TRANSPORT ATP-BINDING PROTEIN CYDC [Bacillus subtilis]
1214	Bt1Gc1558	Bt1G2342	11-388	g1723608	199	163	4.10E-12	38	99	HYPOTHETICAL PROTEIN IN GLVBC 3'REGION [Bacillus subtilis]
1214	Bt1Gc1558	Bt1G2343	446-1285	g1723610	475	562	2.10E-54	42	100	HYPOTHETICAL 31.5 KD PROTEIN IN GLVBC 3'REGION [Bacillus subtilis]
1214	Bt1Gc1558	Bt1G2344	3412-2433	g1172469	299	395	1.10E-36	29	100	PHENYLALANINE-4-HYDROXYLASE (PAH) (PHE-4-MONOOXYGENASE) []
1214	Bt1Gc1558	Bt1G2345	6453-6769	g2497382	528	544	1.70E-52	99	24	TRANSPORASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1215	Bt1Gc1561	Bt1G2346	815-1	g4009481	451	508	1.10E-48	38	77	(AF068902) undecaprenyl-PP-MurNAc-pentapeptide-UDPGlcNAc GlcNAc transferase [Streptococcus pneumoniae]
1215	Bt1Gc1561	Bt1G2347	1927-1222	g3249547	174	272	1.10E-23	29	89	(AF009955) oxidoreductase Rmd [Pseudomonas aeruginosa]
1215	Bt1Gc1561	Bt1G2348	3130-2149	g2127789	620	695	1.70E-68	43	100	capsular polysaccharide biosynthesis protein D homolog - Methanococcus jannaschii [Methanococcus jannaschii]
1215	Bt1Gc1561	Bt1G2349	3531-4184	g2126959	444	433	9.90E-41	44	100	hypothetical protein (212 aa) (ponA operon) - Bacillus subtilis (fragment) [Bacillus subtilis]
1215	Bt1Gc1561	Bt1G2350	5410-4992	g2116755	339	367	9.80E-34	47	99	(D86418) YfmQ [Bacillus subtilis]
1215	Bt1Gc1561	Bt1G2351	6485-5596	g2636549	461	217	9.10E-32	36	100	(Z99124) similar to hypothetical proteins [Bacillus subtilis]
1215	Bt1Gc1561	Bt1G2352	7684-6516	g1881344	593	658	1.40E-64	36	81	(AB001488) SIMILAR TO THE RHIZOPINE CATABOLISM (MOCR) GENE OF RHIZOBIUM MELILOTI. [Bacillus subtilis]
1216	Bt1Gc1564	Bt1G2353	87-257	g1710750	272	204	1.80E-16	98	98	30S RIBOSOMAL PROTEIN S21 (BS-B) [Bacillus subtilis]
1216	Bt1Gc1564	Bt1G2354	276-716	g1730992	477	273	9.00E-24	67	99	HYPOTHETICAL 16.8 KD PROTEIN IN RPSU-PHOH INTERGENIC REGION [Bacillus subtilis]
1216	Bt1Gc1564	Bt1G2355	1633-1914	g3947509	333	274	7.00E-24	68	99	(AJ224829) ORF3 [Bacillus megaterium]
1216	Bt1Gc1564	Bt1G2356	2094-3278	g3947510	1071	956	3.80E-96	50	100	(AJ224829) sporulation protein [Bacillus megaterium]
1216	Bt1Gc1564	Bt1G2357	3297-4268	g3947511	1194	1197	1.10E-121	73	100	(AJ224829) ORF4 [Bacillus megaterium]
1216	Bt1Gc1564	Bt1G2358	4439-6303	g1730999	1488	1355	2.00E-138	46	87	HYPOTHETICAL 79.2 KD PROTEIN IN PHOH-DGKA INTERGENIC REGION [Bacillus subtilis]
1217	Bt1Gc1559	Bt1G2359	1-1476	g1706795	362	278	2.60E-24	33	89	FERRICHROME TRANSPORT PERMEASE PROTEIN FHUB [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1217	Bt1Gc1559	Bt1G2360	482-1483	g1706797	736	526	1.40E-50	42	100	FERRICHRONE TRANSPORT PERMEASE PROTEIN FHUG [Bacillus subtilis]
1217	Bt1Gc1559	Bt1G2361	2173-1526	g120586	315	355	1.80E-32	37	100	L-FUCULOSE PHOSPHATE ALDOLASE [Escherichia coli]
1217	Bt1Gc1559	Bt1G2362	3273-2215	g3183453	817	850	6.40E-85	49	100	HYPOTHETICAL 38.9 KD PROTEIN IN DAT-SPO0E INTERGENIC REGION [Bacillus subtilis]
1217	Bt1Gc1559	Bt1G2363	4524-3295	g2633727	618	726	8.90E-72	37	100	(Z99111) ykrT [Bacillus subtilis]
1217	Bt1Gc1559	Bt1G2364	5104-5599	g1703216	671	673	3.70E-66	77	88	ALKYL HYDROPEROXIDE REDUCTASE C22 PROTEIN (GENERAL STRESS PROTEIN 22) [Bacillus subtilis]
1218	Bt1Gc1562	Bt1G2365	493-1	g1770007	729	562	2.10E-54	86	95	(Z75208) translation initiation factor IF3 [Bacillus subtilis]
1218	Bt1Gc1562	Bt1G2366	2809-876	g135175	2706	2662	6.20E-277	80	100	THREONYL-TRNA SYNTHETASE 1 (THREONINE--TRNA LIGASE) (THRRS) [Bacillus subtilis]
1218	Bt1Gc1562	Bt1G2367	3998-3141	g140017	378	414	1.00E-38	31	100	HYPOTHETICAL 33.3 KD PROTEIN IN DNAI-THRS INTERGENIC REGION (ORF-281) [Bacillus subtilis]
1218	Bt1Gc1562	Bt1G2368	5403-4468	g140025	806	816	2.60E-81	51	100	PRIMOSOMAL PROTEIN DNAI [Bacillus subtilis]
1218	Bt1Gc1562	Bt1G2369	6843-5452	g118711	1038	1122	9.70E-114	47	100	REPLICATION INITIATION AND MEMBRANE ATTACHMENT PROTEIN [Bacillus subtilis]
1218	Bt1Gc1562	Bt1G2370	7460-7005	g2293279	619	542	2.80E-52	75	99	(AF008220) YtcG [Bacillus subtilis]
1219	Bt1Gc1563	Bt1G2371	3333-142	g1769947	5149	5081	0	93	100	(X98455) SNF [Bacillus cereus]
1219	Bt1Gc1563	Bt1G2372	4384-3455	g1769946	1440	1387	8.00E-142	91	100	(X98455) orfI [Bacillus cereus]
1219	Bt1Gc1563	Bt1G2373	5579-4447	g3599667	824	634	5.00E-62	44	100	(AB006424) ybgK [Bacillus subtilis]
1219	Bt1Gc1563	Bt1G2374	7297-5768	g2293209	1859	1645	3.70E-169	70	90	(AF008220) putative malolactic enzyme [Bacillus subtilis]
1220	Bt1Gc1560	Bt1G2375	246-964	g732337	559	447	3.30E-42	50	100	HYPOTHETICAL 25.8 KD PROTEIN IN EPR-GALK INTERGENIC REGION [Bacillus subtilis]
1220	Bt1Gc1560	Bt1G2376	2821-1055	g1945085	715	513	3.30E-49	55	100	(D88802) transmembrane [Bacillus subtilis]
1220	Bt1Gc1560	Bt1G2377	7346-6222	g2226165	782	824	3.70E-82	44	55	(Y14080) hypothetical protein [Bacillus subtilis]
1221	Bt1Gc1565	Bt1G2378	1161-616	g1881232	364	335	2.40E-30	40	95	(AB001488) PROBABLE ACETYLTRANSFERASE. [Bacillus subtilis]
1221	Bt1Gc1565	Bt1G2379	2654-1359	g3183540	1685	1689	8.00E-174	74	100	UDP-N-ACETYLMURAMATE--ALANINE LIGASE (UDP-N-ACETYLMURANOYL-L-ALANINE SYNTHETASE) [Bacillus subtilis]
1221	Bt1Gc1565	Bt1G2380	6687-4105	g2293215	1750	1790	1.60E-184	54	100	(AF008220) YtpT [Bacillus subtilis]
1221	Bt1Gc1565	Bt1G2381	7425-5506	g1076839	231	87	0.078	27	40	protein kinase - slime mold (Dictyostelium discoideum) [Dictyostelium discoideum]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1222	Bt1Gc1566	Bt1G2382	81-791	g1169717	699	714	1.70E-70	56	100	ANAEROBIC REGULATORY PROTEIN [Bacillus subtilis]
1222	Bt1Gc1566	Bt1G2383	7220-3656	g2117582	5360	5206	0	81	97	respiratory nitrate reductase alpha chain - Bacillus subtilis [Bacillus subtilis]
1223	Bt1Gc1569	Bt1G2384	1-2186	g2633696	938	1071	2.50E-108	39	77	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
1223	Bt1Gc1569	Bt1G2385	2348-2205	g2127281	208	208	6.90E-17	83	98	hypothetical 5.3K protein - Bacillus thuringiensis [Bacillus thuringiensis]
1223	Bt1Gc1569	Bt1G2386	3295-2441	g2127289	1255	1255	7.80E-128	83	100	transcriptional activator plcR - Bacillus thuringiensis [Bacillus thuringiensis]
1223	Bt1Gc1569	Bt1G2387	2264-5170	g128529	220	423	1.10E-39	33	53	BACILLOLYSIN PRECURSOR (NEUTRAL PROTEASE) [Paenibacillus polymyxa]
1224	Bt1Gc1568	Bt1G2388	5658-5278	g732333	339	351	4.90E-32	52	99	HYPOTHETICAL 14.4 KD PROTEIN IN EPR-GALK INTERGENIC REGION [Bacillus subtilis]
1224	Bt1Gc1568	Bt1G2389	6293-9003	g2506893	55	175	5.50E-12	27	18	HYPOTHETICAL PROTEIN KIAA0083 [Homo sapiens]
1224	Bt1Gc1568	Bt1G2390	8358-9549	g2894215	545	554	1.50E-53	35	100	(AL021841) amIB [Mycobacterium tuberculosis]
1224	Bt1Gc1568	Bt1G2391	10146-9892	g2632968	178	179	8.20E-14	35	99	(Z99107) similar to transcriptional regulator (Lrp/AsnC family) [Bacillus subtilis]
1224	Bt1Gc1568	Bt1G2392	10332-10964	g4240421	463	476	2.80E-45	47	82	(AF080235) reductase homolog [Streptomyces cyanogenus]
1225	Bt1Gc1567	Bt1G2393	1-869	g730103	848	797	2.70E-79	57	85	NARA PROTEIN [Bacillus subtilis]
1225	Bt1Gc1567	Bt1G2394	890-1903	g2633798	927	866	1.30E-86	51	100	(Z99111) molybdopterin biosynthesis protein [Bacillus subtilis]
1225	Bt1Gc1567	Bt1G2395	1947-3236	g2633799	1328	1332	5.40E-136	61	100	(Z99111) molybdopterin biosynthesis protein [Bacillus subtilis]
1225	Bt1Gc1567	Bt1G2396	3287-3762	g2633801	423	456	3.60E-43	54	99	(Z99111) molybdopterin converting factor (subunit 2) [Bacillus subtilis]
1225	Bt1Gc1567	Bt1G2397	3700-3930	g2633802	184	184	2.40E-14	47	99	(Z99111) molybdopterin converting factor (subunit 1) [Bacillus subtilis]
1225	Bt1Gc1567	Bt1G2398	4011-5195	g1171655	1391	1001	6.40E-101	65	100	NITRITE EXTRUSION PROTEIN (NITRITE FACILITATOR) [Bacillus subtilis]
1225	Bt1Gc1567	Bt1G2399	6528-6055	g2462962	224	218	6.00E-18	33	99	(AJ000974) YlnF protein [Bacillus subtilis]
1225	Bt1Gc1567	Bt1G2400	7243-5768	g2462961	265	310	1.10E-27	30	100	(AJ000974) YlnE protein [Bacillus subtilis]
1225	Bt1Gc1567	Bt1G2401	8467-7948	g267052	479	414	1.00E-38	57	73	UROPORPHYRIN-III C-METHYLTRANSFERASE (UROGEN III METHYLASE) (SUMT) (UROPORPHYRINOGEN III METHYLASE) (UROM) [Bacillus megaterium]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1226	Bt1Gc1570	Bt1G2402	303-1016	g4481749	918	922	1.50E-92	74	100	(AF007865) BacR [Bacillus licheniformis]
1226	Bt1Gc1570	Bt1G2403	997-2031	g4481748	771	761	1.70E-75	52	100	(AF007865) BacS [Bacillus licheniformis]
1226	Bt1Gc1570	Bt1G2404	2097-3014	g1168621	1276	1231	2.70E-125	79	100	BACITRACIN TRANSPORT ATP-BINDING PROTEIN BCRA [Bacillus licheniformis]
1226	Bt1Gc1570	Bt1G2405	3121-3744	g1168622	611	492	5.60E-47	54	100	BACITRACIN TRANSPORT PERMEASE PROTEIN BCRB [Bacillus licheniformis]
1226	Bt1Gc1570	Bt1G2406	3756-4556	g1705428	713	561	2.70E-54	55	100	BACITRACIN RESISTANCE PROTEIN (PUTATIVE UNDECAPRENOL KINASE) [Escherichia coli]
1227	Bt1Gc1572	Bt1G2407	281-1	g1881228	159	173	2.40E-12	36	20	(AB001488) SIMILAR TO ENZYMES WHICH ACT VIA AN ATP-DEPENDENT COVALENT BINDING OF AMP TO THEIR SUBSTRATE. [Bacillus subtilis]
1227	Bt1Gc1572	Bt1G2408	958-17	g3183273	412	508	1.10E-48	48	99	PUTATIVE NAD(P)H OXIDOREDUCTASE YDEQ [Bacillus subtilis]
1227	Bt1Gc1572	Bt1G2409	1070-1448	g2293276	297	340	7.10E-31	52	99	(AF008220) YtcD [Bacillus subtilis]
1227	Bt1Gc1572	Bt1G2410	2134-1841	g4894353	142	173	3.50E-13	37	99	(AF065404) pXO1-138 [Bacillus anthracis]
1227	Bt1Gc1572	Bt1G2411	6788-3625	g3721552	209	321	5.40E-28	34	35	(AB010958) xylanase [Clostridium thermocellum]
1227	Bt1Gc1572	Bt1G2412	5640-4477	g2635915	265	367	9.80E-34	25	100	(Z99121) similar to hypothetical proteins [Bacillus subtilis]
1227	Bt1Gc1572	Bt1G2413	7319-6021	g2633717	321	296	3.30E-26	25	100	(Z99111) ykrl [Bacillus subtilis]
1227	Bt1Gc1572	Bt1G2414	8043-7318	g2633716	565	587	4.80E-57	49	100	(Z99111) similar to RNA polymerase sigma factor [Bacillus subtilis]
1227	Bt1Gc1572	Bt1G2415	9330-6993	g1169894	77	116	3.40E-06	34	19	MINOR TEICHOIC ACIDS BIOSYNTHESIS PROTEIN GGAA [Bacillus subtilis]
1228	Bt1Gc1574	Bt1G2416	142-405	g1710749	250	254	9.20E-22	58	99	30S RIBOSOMAL PROTEIN S20 (BS20) [Bacillus subtilis]
1228	Bt1Gc1574	Bt1G2417	4636-1999	g1303798	1185	1229	4.40E-125	36	100	(D84432) ComEC [Bacillus subtilis]
1228	Bt1Gc1574	Bt1G2418	4899-4331	g418361	641	661	6.90E-65	65	99	COME OPERON PROTEIN 2 [Bacillus subtilis]
1228	Bt1Gc1574	Bt1G2419	5562-4966	g729154	320	365	1.60E-33	40	100	COME OPERON PROTEIN 1 [Bacillus subtilis]
1228	Bt1Gc1574	Bt1G2420	5640-6458	g1303795	815	763	1.10E-75	55	100	(D84432) ComED [Bacillus subtilis]
1228	Bt1Gc1574	Bt1G2421	7389-6631	g1730986	541	598	3.30E-58	48	100	HYPOTHETICAL 28.3 KD PROTEIN IN AROD-COMER INTERGENIC REGION [Bacillus subtilis]
1228	Bt1Gc1574	Bt1G2422	7742-7389	g1730985	436	368	7.70E-34	69	99	HYPOTHETICAL 13.3 KD PROTEIN IN AROD-COMER INTERGENIC REGION [Bacillus subtilis]
1228	Bt1Gc1574	Bt1G2423	8308-7751	g1730984	609	611	1.40E-59	63	99	HYPOTHETICAL 21.3 KD PROTEIN IN AROD-COMER INTERGENIC REGION [Bacillus subtilis]
1228	Bt1Gc1574	Bt1G2424	8864-8301	g3323048	196	239	3.60E-20	29	100	(AE001246) conserved hypothetical protein [Treponema

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1228	Bt1Gc1574	Bt1G2425	9298-9011	g1730982	397	397	6.50E-37	79	99	pallidum] HYPOTHETICAL 10.8 KD PROTEIN IN AROD-COMER INTERGENIC REGION [Bacillus subtilis]
1228	Bt1Gc1574	Bt1G2426	9523-9298	g1703404	100	145	8.20E-10	32	26	SHIKIMATE 5-DEHYDROGENASE [Bacillus subtilis]
1229	Bt1Gc1573	Bt1G2427	940-563	g1781143	172	145	3.30E-10	33	99	(Z83866) hypothetical protein Rv3070 [Mycobacterium tuberculosis]
1229	Bt1Gc1573	Bt1G2428	1320-949	g2983116	158	101	1.50E-05	32	99	(AE000690) hypothetical protein [Aquifex aeolicus]
1229	Bt1Gc1573	Bt1G2429	3312-4264	g4512389	286	247	5.10E-20	26	48	(AB011838) methyl-accepting chemotaxis protein [Bacillus halodurans]
1229	Bt1Gc1573	Bt1G2430	4467-4838	g2447088	92	142	6.80E-10	30	99	(U42580) A612L [Paramecium bursaria Chlorella virus 1]
1229	Bt1Gc1573	Bt1G2431	4384-6159	g1653724	442	531	4.10E-51	38	100	(D90916) hypothetical protein [Synecocystis sp.]
1229	Bt1Gc1573	Bt1G2432	6532-7199	g494692	680	699	6.50E-69	58	70	Tyrosyl-Transfer Rna Synthetase (E.C.6.1.1.1) Mutant With Thr 51 Replaced By Pro (T51p) []
1230	Bt1Gc1571	Bt1G2433	47-751	g1945655	635	661	6.90E-65	54	100	(Z94043) hypothetical protein [Bacillus subtilis]
1230	Bt1Gc1571	Bt1G2434	747-1808	g1945656	540	496	2.10E-47	35	100	(Z94043) hypothetical protein [Bacillus subtilis]
1230	Bt1Gc1571	Bt1G2435	2671-4596	g1945658	950	841	5.80E-84	33	100	(Z94043) hypothetical protein [Bacillus subtilis]
1230	Bt1Gc1571	Bt1G2436	4643-6562	g1945658	702	410	2.50E-52	29	100	(Z94043) hypothetical protein [Bacillus subtilis]
1230	Bt1Gc1571	Bt1G2437	6656-7501	g130048	836	825	2.90E-82	59	100	PREPHENATE DEHYDRATASE (PDT) [Bacillus subtilis]
1230	Bt1Gc1571	Bt1G2438	8266-7613	g2443233	428	460	1.40E-43	40	100	(D86417) YfIk [Bacillus subtilis]
1230	Bt1Gc1571	Bt1G2439	8874-8344	g732258	521	536	1.20E-51	56	99	HYPOTHETICAL 19.7 KD PROTEIN IN PHEA-NIFS INTERGENIC REGION (ORF1) [Bacillus subtilis]
1231	Bt1Gc1578	Bt1G2440	1807-2193	g80103	292	311	8.40E-28	41	99	beta-lactamase repressor protein - Bacillus licheniformis [Bacillus licheniformis]
1231	Bt1Gc1578	Bt1G2441	2241-5046	g2791915	147	313	2.50E-27	23	55	(Y13095) MccR1 protein [Staphylococcus sciuri]
1231	Bt1Gc1578	Bt1G2442	4320-4892	g418449	625	539	5.80E-52	63	99	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN HEMY-GLTT INTERGENIC REGION (ORFA) [Bacillus subtilis]
1231	Bt1Gc1578	Bt1G2443	5118-8432	g2851530	1363	1362	3.60E-139	38	100	HYPOTHETICAL 84.1 KD PROTEIN IN HEMY-GLTT INTERGENIC REGION (ORFB) [Bacillus subtilis]
1232	Bt1Gc1575	Bt1G2444	112-1611	g4514332	2194	2155	3.30E-223	83	100	(AB013369) unknown [Bacillus halodurans]
1232	Bt1Gc1575	Bt1G2445	2687-1665	g1787795	1031	1076	7.20E-109	58	100	(AE000249) putative LACI-type transcriptional regulator [Escherichia coli]
1232	Bt1Gc1575	Bt1G2446	3694-2705	g3025114	671	483	5.00E-46	41	100	HYPOTHETICAL ABC TRANSPORTER PERMEASE PROTEIN YDEZ [Escherichia coli]
1232	Bt1Gc1575	Bt1G2447	6137-4629	g3025112	894	908	4.60E-91	37	100	HYPOTHETICAL ABC TRANSPORTER A TP-BINDING

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1232	Bt1Gc1575	Bt1G2448	6296-7237	g3025110	455	520	6.00E-50	38	100	PROTEIN YDEX [Escherichia coli] HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN HIPB-UXAB INTERGENIC REGION [Escherichia coli]
1232	Bt1Gc1575	Bt1G2449	6401-8806	g2493485	1432	1394	1.50E-142	54	100	HYPOTHETICAL SUGAR KINASE IN HIPB-UXAB INTERGENIC REGION [Escherichia coli]
1232	Bt1Gc1575	Bt1G2450	9114-9986	g2495606	507	586	6.10E-57	41	100	HYPOTHETICAL 31.9 KD PROTEIN IN HIPB-UXAB INTERGENIC REGION [Escherichia coli]
1233	Bt1Gc1576	Bt1G2451	1-721	g1724006	827	663	4.20E-65	69	52	HYPOTHETICAL 49.0 KD PROTEIN IN CSPB-GLPP INTERGENIC REGION [Bacillus subtilis]
1233	Bt1Gc1576	Bt1G2452	4649-3587	g4009481	608	604	7.50E-59	40	100	(AF068902) undecaprenyl-PP-MurNAc-pentapeptide-UDPGlcNAc GlcNAc transferase [Streptococcus pneumoniae]
1233	Bt1Gc1576	Bt1G2453	7301-4674	g3122390	1258	1110	4.40E-143	37	78	5-METHYLTETRAHYDROFOLATE--HOMOCYSTEINE METHYLTRANSFERASE (METHIONINE SYNTHASE, VITAMIN-B12 DEPENDENT ISOZYME) (MS) [Mycobacterium tuberculosis]
1234	Bt1Gc1577	Bt1G2454	1-1185	g1169253	1085	1021	4.90E-103	57	90	ANAEROBIC C4-DICARBOXYLATE TRANSPORTER DCUB [Haemophilus influenzae Rd]
1234	Bt1Gc1577	Bt1G2455	1962-1297	g732341	665	592	1.40E-57	59	100	THIAMIN-PHOSPHATE PYROPHOSPHORYLASE (TMP-PYROPHOSPHORYLASE) (TMP-PPASE) (THIAMIN-PHOSPHATE SYNTHASE) [Bacillus subtilis]
1234	Bt1Gc1577	Bt1G2456	2787-1975	g732340	804	819	1.20E-81	60	100	HYDROXYETHYLTHIAZOLE KINASE (4-METHYL-5-BETA-HYDROXYETHYLTHIAZOLE KINASE) (THZ KINASE) (TH KINASE) [Bacillus subtilis]
1234	Bt1Gc1577	Bt1G2457	3199-4691	g1881236	1658	1645	3.70E-169	64	68	(AB001488) PROBABLE DNA TOPOISOMERASE III [Bacillus subtilis]
1235	Bt1Gc1580	Bt1G2458	1001-1	g2293322	888	833	4.10E-83	54	75	(AF008220) branch-chain amino acid transporter [Bacillus subtilis]
1235	Bt1Gc1580	Bt1G2459	2188-323	g418534	554	603	9.60E-59	48	89	HYPOTHETICAL 32.5 KD PROTEIN IN PEPE-LYSC INTERGENIC REGION [Escherichia coli]
1235	Bt1Gc1580	Bt1G2460	4726-2369	g3914084	2531	2282	1.20E-236	63	100	MUTS2 PROTEIN [Bacillus subtilis]
1235	Bt1Gc1580	Bt1G2461	6458-4749	g1770035	1974	1733	1.70E-178	65	100	(Z75208) hypothetical protein [Bacillus subtilis]
1235	Bt1Gc1580	Bt1G2462	7059-6526	g1770034	387	352	3.80E-32	45	99	(Z75208) hypothetical protein [Bacillus subtilis]
1235	Bt1Gc1580	Bt1G2463	7324-7070	g1770033	269	236	7.50E-20	58	99	(Z75208) hypothetical protein [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1236	Bt1Gc1581	Bt1G2464	1532-1	g2632162	106	251	1.90E-21	32	57	(AJ001118) monoglyceride lipase [Mus musculus]
1236	Bt1Gc1581	Bt1G2465	748-1371	g259337	379	414	1.00E-38	40	100	(AJ224767) hypothetical protein [Acinetobacter sp. ADP1]
1236	Bt1Gc1581	Bt1G2466	2808-1585	g140388	930	835	2.50E-83	45	100	HYPOTHETICAL 43.9 KD PROTEIN IN MSYB-HTRB INTERGENIC REGION (ORF1) [Escherichia coli]
1236	Bt1Gc1581	Bt1G2467	3824-2955	g4584123	1109	1125	4.60E-114	75	100	(AJ007788) hypothetical protein [Bacillus cereus]
1236	Bt1Gc1581	Bt1G2468	4547-3834	g4584122	1109	784	6.30E-78	95	100	(AJ007788) hypothetical protein [Bacillus cereus]
1236	Bt1Gc1581	Bt1G2469	4752-5852	g4584121	1827	1776	4.80E-183	95	100	(AJ007788) related sequence M24150 [Bacillus cereus]
1236	Bt1Gc1581	Bt1G2470	7478-6219	g4584120	2130	2070	3.40E-214	98	100	(AJ007788) threonine dehydratase [Bacillus cereus]
1236	Bt1Gc1581	Bt1G2471	9189-7525	g1708465	2074	1869	6.70E-193	69	100	DIHYDROXY-ACID DEHYDRATASE (DAD) (VEGETATIVE PROTEIN 110) (VEG110) [Bacillus subtilis]
1237	Bt1Gc1582	Bt1G2472	2424-1024	g417115	1366	1193	2.90E-121	54	100	PROTOPORPHYRINOGEN OXIDASE (PPO) [Bacillus subtilis]
1237	Bt1Gc1582	Bt1G2473	3373-2444	g417116	1204	1211	3.60E-123	72	100	FERROCHELATASE (PROTOHEME FERRO-LYASE) (HEME SYNTHETASE) [Bacillus subtilis]
1237	Bt1Gc1582	Bt1G2474	4437-3379	g416890	1382	1402	2.10E-143	72	100	UROPORPHYRINOGEN DECARBOXYLASE (UPD) [Bacillus subtilis]
1237	Bt1Gc1582	Bt1G2475	6938-4597	g585645	1806	1854	2.60E-191	49	100	PENICILLIN-BINDING PROTEIN 1A (PBP-1A) [Bacillus subtilis]
1237	Bt1Gc1582	Bt1G2476	6771-7354	g586801	340	425	7.00E-40	50	99	HYPOTHETICAL 18.8 KD PROTEIN IN ECSC-PBPF INTERGENIC REGION [Bacillus subtilis]
1237	Bt1Gc1582	Bt1G2477	8967-7549	g1881270	468	527	1.10E-50	30	100	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
1238	Bt1Gc1579	Bt1G2478	831-1	g585159	1031	937	3.90E-94	74	43	CELL DIVISION PROTEIN FTSH HOMOLOG [Bacillus subtilis]
1238	Bt1Gc1579	Bt1G2479	1459-920	g585264	745	745	8.60E-74	79	99	HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE (HGPRT) (HGPRTASE) [Bacillus subtilis]
1238	Bt1Gc1579	Bt1G2480	2913-1459	g586892	1052	1111	1.40E-112	45	100	HYPOTHETICAL 55.1 KD PROTEIN IN SPOIIE-HPT INTERGENIC REGION [Bacillus subtilis]
1238	Bt1Gc1579	Bt1G2481	5588-3108	g586018	2291	1654	1.10E-209	55	100	STAGE II SPORULATION PROTEIN E [Bacillus subtilis]
1238	Bt1Gc1579	Bt1G2482	6649-6167	g586889	523	558	5.60E-54	88	99	HYPOTHETICAL 14.2 KD PROTEIN IN DIVIC-SPOIIE INTERGENIC REGION [Bacillus subtilis]
1238	Bt1Gc1579	Bt1G2483	7105-6731	g585054	217	203	2.30E-16	41	99	CELL DIVISION PROTEIN DIVIC [Bacillus subtilis]
1238	Bt1Gc1579	Bt1G2484	7746-7114	g586888	412	384	1.50E-35	45	100	HYPOTHETICAL 25.1 KD PROTEIN IN MFD-DIVIC INTERGENIC REGION [Bacillus subtilis]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1238	Bt1Gc1579	Bt1G2485	8051-7746	g586887	339	301	9.70E-27	67	99	HYPOTHETICAL 11.4 KD PROTEIN IN MFD-DIVIC INTERGENIC REGION [Bacillus subtilis]
1238	Bt1Gc1579	Bt1G2486	8385-8128	g586886	344	347	1.30E-31	79	99	HYPOTHETICAL 9.7 KD PROTEIN IN MFD-DIVIC INTERGENIC REGION [Bacillus subtilis]
1238	Bt1Gc1579	Bt1G2487	8718-8397	g586885	250	291	3.10E-25	52	22	HYPOTHETICAL 56.1 KD PROTEIN IN MFD-DIVIC INTERGENIC REGION [Bacillus subtilis]
1239	Bt1Gc1583	Bt1G2488	659-1351	g3183122	853	806	3.00E-80	70	100	5'-METHYLTHIOADENOSINE / S-ADENOSYLHOMOCYSTEINE NUCLEOSIDASE (MTA/SAH NUCLEOSIDASE) (P46) [Bacillus subtilis]
1239	Bt1Gc1583	Bt1G2489	1437-2357	g1934605	1005	946	4.30E-95	62	100	(U93874) cysteine synthase [Bacillus subtilis]
1239	Bt1Gc1583	Bt1G2490	2364-3500	g1934606	1341	1307	2.40E-133	68	100	(U93874) cystathionine gamma-lyase [Bacillus subtilis]
1239	Bt1Gc1583	Bt1G2491	3777-6365	g2492737	2459	2347	1.50E-243	56	100	ALCOHOL DEHYDROGENASE 2 (ADH) / ALCTALDEHYDE DEHYDROGENASE (ACDH) [Entamoeba histolytica]
1240	Bt1Gc1586	Bt1G2492	1-556	g2293243	649	671	6.00E-66	67	40	(AF008220) arginine succinate lyase [Bacillus subtilis]
1240	Bt1Gc1586	Bt1G2493	2521-2099	g2622094	162	191	4.40E-15	32	99	(AE000872) conserved protein [Methanobacterium thermoautotrophicum]
1240	Bt1Gc1586	Bt1G2494	2711-3472	g2293244	919	932	1.30E-93	68	100	(AF008220) 3-ketoacyl-ACP reductase [Bacillus subtilis]
1240	Bt1Gc1586	Bt1G2495	3685-4815	g4803749	1419	1442	1.20E-147	74	100	(AB013821) alanine dehydrogenase [Enterobacter aerogenes]
1240	Bt1Gc1586	Bt1G2496	5422-5760	g1881251	376	390	3.60E-36	59	99	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
1240	Bt1Gc1586	Bt1G2497	5745-6527	g1881342	392	441	1.40E-41	34	100	(AB001488) SIMILAR TO BACILLUS CEREUS ZWITTERMICIN A-RESISTANCE GENE. [Bacillus subtilis]
1241	Bt1Gc1585	Bt1G2498	1302-482	g120474	815	829	1.10E-82	58	100	FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (FAPY-DNA GLYCOSYLASE) [Bacillus firmus]
1241	Bt1Gc1585	Bt1G2499	3954-1318	g2293272	3115	2861	5.10E-298	69	100	(AF008220) DNA-polymerase I [Bacillus subtilis]
1241	Bt1Gc1585	Bt1G2500	5997-4234	g130130	1344	1350	6.70E-138	51	100	ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN PHOR [Bacillus subtilis]
1241	Bt1Gc1585	Bt1G2501	6709-5993	g2293270	906	911	2.20E-91	74	100	(AF008220) signal transduction regulator [Bacillus subtilis]
1242	Bt1Gc1584	Bt1G2502	1-1246	g135736	418	656	2.30E-64	38	72	THERMOLYSIN PRECURSOR (THERMOSTABLE NEUTRAL PROTEINASE) [Bacillus stearothermophilus]
1242	Bt1Gc1584	Bt1G2503	3715-5196	g1730175	1200	1244	1.10E-126	48	100	GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE (G6PD) (VEGETATIVE PROTEIN 11) (VEG11) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1242	Bt1Gc1584	Bt1G2504	5257-7245	g2507484	2003	1771	1.60E-182	57	100	TRANSKETOLASE [Bacillus subtilis]
1243	Bt1Gc1591	Bt1G2505	1475-404	g728897	1410	1189	7.70E-121	77	100	PHOSPHO-2-DEHYDRO-3-DEOXYHEPTONATE ALDOLASE (PHOSPHO-2-KETO-3-DEOXYHEPTONATE ALDOLASE) (DAHP SYNTHETASE) (3-DEOXY-D-ARABINO-HEPTULOSONATE 7-PHOSPHATE SYNTHASE) / CHORISMATE MUTASE [Bacillus subtilis]
1243	Bt1Gc1591	Bt1G2506	3106-2783	g732304	268	255	7.20E-22	44	99	HYPOTHETICAL 12.4 KD PROTEIN IN MURC-AROA INTERGENIC REGION (ORF2) (ORF3) [Bacillus subtilis]
1243	Bt1Gc1591	Bt1G2507	3618-3199	g732301	355	244	1.10E-20	50	99	HYPOTHETICAL 15.7 KD PROTEIN IN MURC-AROA INTERGENIC REGION (ORF1) [Bacillus subtilis]
1243	Bt1Gc1591	Bt1G2508	4439-5326	g2507254	83	208	2.70E-16	39	30	AMINOPEPTIDASE AMPS [Bacillus subtilis]
1244	Bt1Gc1588	Bt1G2509	Jan-79	g4098489	541	567	6.30E-55	53	56	(U78600) putative ptsG protein [Streptococcus mutans]
1244	Bt1Gc1588	Bt1G2510	1675-4626	g538903	997	1208	7.40E-123	32	100	phage infection protein precursor - Lactococcus lactis subsp. lactis (strain C2) []
1244	Bt1Gc1588	Bt1G2511	5797-4670	g1881236	1183	1161	7.10E-118	63	52	(AB001488) PROBABLE DNA TOPOISOMERASE III [Bacillus subtilis]
1245	Bt1Gc1589	Bt1G2512	582-1	g586858	659	626	3.50E-61	68	98	HYPOTHETICAL 21.4 KD PROTEIN IN DACA-SERS INTERGENIC REGION [Bacillus subtilis]
1245	Bt1Gc1589	Bt1G2513	1485-604	g586857	1291	1232	2.10E-125	86	100	31.5 KD GUANYLYLATED PROTEIN IN DACA-SERS INTERGENIC REGION (SUPEROXIDE-INDUCIBLE PROTEIN 7) (SOI7) [Bacillus subtilis]
1245	Bt1Gc1589	Bt1G2514	2965-1652	g585034	1165	1162	5.60E-118	54	100	D-ALANYL-D-ALANINE CARBOXYPEPTIDASE PRECURSOR (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) (CPASE) (PBP5) [Bacillus subtilis]
1245	Bt1Gc1589	Bt1G2515	4533-3070	g2127114	2029	2033	2.80E-210	79	100	IMP dehydrogenase - Bacillus subtilis [Bacillus subtilis]
1245	Bt1Gc1589	Bt1G2516	4655-5626	g586856	652	651	7.90E-64	45	100	HYPOTHETICAL 38.0 KD PROTEIN IN GIRA-GUAB INTERGENIC REGION [Bacillus subtilis]
1246	Bt1Gc1587	Bt1G2517	1-366	g141471	464	417	4.90E-39	72	65	TRANSALDOLASE-LIKE PROTEIN (20 KD PHOSPHOPROTEIN ORFU) (CSI9) [Bacillus subtilis]
1246	Bt1Gc1587	Bt1G2518	1176-3239	g124464	2632	2519	8.90E-262	70	100	IMMUNE INHIBITOR A PRECURSOR [Bacillus thuringiensis]
1246	Bt1Gc1587	Bt1G2519	5444-4022	g461637	168	276	4.30E-24	23	100	MULTIDRUG RESISTANCE PROTEIN 1 (MULTIDRUG-EFFLUX TRANSPORTER 1) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1246	Bt1Gc1587	Bt1G2520	5456-4265	g728970	180	293	6.80E-26	21	100	MULTIDRUG RESISTANCE PROTEIN 2 (MULTIDRUG-EFFLUX TRANSPORTER 2) [Bacillus subtilis]
1246	Bt1Gc1587	Bt1G2521	5650-6699	g2522016	1099	1107	3.80E-112	64	100	(AB007638) dehydrogenase [Bacillus subtilis]
1247	Bt1Gc1594	Bt1G2522	1720-425	g2851670	1340	1190	6.00E-121	59	100	HYPOTHETICAL 49.0 KD PROTEIN IN BLTD-TRKA INTERGENIC REGION [Bacillus subtilis]
1247	Bt1Gc1594	Bt1G2523	2274-1840	g2226212	204	194	2.10E-15	36	99	(Y14082) hypothetical protein [Bacillus subtilis]
1247	Bt1Gc1594	Bt1G2524	3387-4166	g1197640	876	893	1.80E-89	61	100	(U46859) DdhA [Yersinia enterocolitica (type 0:8)]
1247	Bt1Gc1594	Bt1G2525	4169-5224	g421277	843	901	2.50E-90	44	100	CDP-glucose-4,6-dehydratase - Yersinia pseudotuberculosis [Yersinia pseudotuberculosis]
1247	Bt1Gc1594	Bt1G2526	5161-6451	g1651978	910	903	1.60E-90	45	100	(D90901) hypothetical protein [Synechocystis sp.]
1248	Bt1Gc1590	Bt1G2527	1475-159	g585605	1420	1365	1.70E-139	71	100	LIPOAMIDE ACYLTRANSFERASE COMPONENT OF BRANCHED-CHAIN ALPHA-KETO ACID DEHYDROGENASE COMPLEX (E2) (DIHYDROLIPOAMIDE BRANCHED CHAIN TRANSACYLASE) [Bacillus subtilis]
1248	Bt1Gc1590	Bt1G2528	2474-1494	g585607	1540	1540	4.90E-158	89	100	2-OXOISOVALERATE DEHYDROGENASE BETA SUBUNIT (BRANCHED-CHAIN ALPHA-KETO ACID DEHYDROGENASE COMPONENT BETA CHAIN (E1)) (BCKDH E1-BETA) [Bacillus subtilis]
1248	Bt1Gc1590	Bt1G2529	3480-2491	g585606	1423	1286	4.00E-131	80	100	2-OXOISOVALERATE DEHYDROGENASE ALPHA SUBUNIT (BRANCHED-CHAIN ALPHA-KETO ACID DEHYDROGENASE COMPONENT ALPHA CHAIN (E1)) (BCKDH E1-ALPHA) [Bacillus subtilis]
1248	Bt1Gc1590	Bt1G2530	4938-3520	g1706441	1795	1743	1.50E-179	73	100	LIPOAMIDE DEHYDROGENASE COMPONENT OF BRANCHED-CHAIN ALPHA-KETO ACID DEHYDROGENASE COMPLEX (E3) (DIHYDROLIPOAMIDE DEHYDROGENASE) (LPD-VAL) [I]
1248	Bt1Gc1590	Bt1G2531	7234-6137	g1256412	1868	1720	4.10E-177	100	100	(U51099) leucine dehydrogenase LeuDH [Bacillus cereus]
1248	Bt1Gc1590	Bt1G2532	8168-7263	g1709891	819	788	2.40E-78	59	100	PROBABLE PHOSPHATE BUTYRYLTRANSFERASE (PHOSPHOTRANSBUTYRYLASE) [Bacillus subtilis]
1248	Bt1Gc1590	Bt1G2533	10381-8367	g1731060	2059	2038	8.30E-211	61	97	PUTATIVE SIGMA L-DEPENDENT TRANSCRIPTIONAL REGULATOR IN MMGE-BFMBA INTERGENIC REGION [Bacillus subtilis]
1249	Bt1Gc1592	Bt1G2534	1-1080	g584749	1048	883	8.90E-98	57	98	N-ACYL-L-AMINO ACID AMIDOHYDROLASE

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1249	Bt1Gc1592	Bt1G2535	1349-2815	g2226174	1780	1416	6.80E-145	67	100	(AMINOACYLASE) [Bacillus stearothermophilus]
1249	Bt1Gc1592	Bt1G2536	2372-4512	g2634159	992	935	6.30E-94	40	100	(Y14081) hypothetical protein [Bacillus subtilis]
1249	Bt1Gc1592	Bt1G2537	4534-5631	g2634160	350	198	4.30E-14	25	100	(Z99113) similar to spore germination protein [Bacillus subtilis]
1249	Bt1Gc1592	Bt1G2538	5624-6160	g4514336	200	234	1.20E-19	32	99	(Z99113) similar to spore germination protein [Bacillus subtilis]
1249	Bt1Gc1592	Bt1G2539	8821-7020	g2498377	352	517	1.20E-49	29	85	(AB013370) YndF1 [Bacillus halodurans]
1249	Bt1Gc1592	Bt1G2540	8821-6209	g2621259	329	494	3.40E-47	32	59	FERROUS IRON TRANSPORT PROTEIN B HOMOLOG [Methanococcus jannaschii]
1250	Bt1Gc1593	Bt1G2541	2547-880	g3123077	260	351	4.90E-32	34	100	(AE000808) ferrous iron transport protein B [Methanobacterium thermoautotrophicum]
1250	Bt1Gc1593	Bt1G2542	3966-2215	g3256824	279	302	7.60E-27	25	100	PUTATIVE GLYCOSYL TRANSFERASE HI0868 [Haemophilus influenzae Rd]
1250	Bt1Gc1593	Bt1G2543	4659-4121	g2454556	452	509	8.80E-49	53	99	(AP000002) 432aa long hypothetical lipopolysaccharide O-side chain biosynthesis protein (O-antigen transporter) [Pyrococcus horikoshii]
1250	Bt1Gc1593	Bt1G2544	5552-4680	g585225	1195	1149	1.30E-116	80	100	UTP--GLUCOSE-1-PHOSPHATE URIDYLTRANSFERASE (UDP-GLUCOSE PYROPHOSPHORYLASE) (ALPHA-D-GLUCOSYL-1-PHOSPHATE URIDYLTRANSFERASE) (URIDINE DIPHOSPHOGLUCOSE PYROPHOSPHORYLASE) (GENERAL STRESS PROTEIN 33) [Bacillus subtilis]
1250	Bt1Gc1593	Bt1G2545	6565-5801	g1894743	728	671	6.00E-66	54	100	(Z92952) product similar to Staphylococcus aureus CapC protein [Bacillus subtilis]
1250	Bt1Gc1593	Bt1G2546	7369-6664	g1894742	615	593	1.10E-57	55	100	(Z92952) product similar to Staphylococcus aureus CapB protein [Bacillus subtilis]
1250	Bt1Gc1593	Bt1G2547	8111-7371	g1894741	614	597	4.20E-58	50	100	(Z92952) product similar to Staphylococcus aureus CapA protein [Bacillus subtilis]
1251	Bt1Gc1595	Bt1G2548	1-932	g2226226	954	552	5.50E-67	61	67	(Y14083) Hypothetical protein [Bacillus subtilis]
1251	Bt1Gc1595	Bt1G2549	2098-1469	g98423	697	617	3.20E-60	62	100	orotate phosphoribosyltransferase (EC 2.4.2.10) - Bacillus subtilis []
1251	Bt1Gc1595	Bt1G2550	3725-2787	g1172785	1320	1332	5.40E-136	80	100	DIHYDROOROTATE DEHYDROGENASE

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1251	Bt1Gc1595	Bt1G2551	4501-3725	g1176556	837	782	1.00E-77	60	100	(DIHYDROOROTATE OXIDASE) (DHODEHASE) [Bacillus caldolyticus]
1251	Bt1Gc1595	Bt1G2552	7610-4504	g115626	4086	4004	0	77	97	HYPOTHETICAL 27.6 KD PROTEIN IN PYRAB-PYRD INTERGENIC REGION (ORF2) [Bacillus caldolyticus] CARBAMOYL-PHOSPHATE SYNTHASE, PYRIMIDINE-SPECIFIC, LARGE CHAIN (CARBAMOYL-PHOSPHATE SYNTHETASE AMMONIA CHAIN) [Bacillus subtilis]
1252	Bt1Gc1598	Bt1G2553	2640-683	g1176954	688	735	9.90E-73	56	74	HYPOTHETICAL 36.9 KD PROTEIN IN ACDA 5'REGION [Bacillus subtilis]
1252	Bt1Gc1598	Bt1G2554	1472-2907	g136144	2449	2407	6.60E-250	99	100	TRANSPPOSASE FOR INSERTION SEQUENCE ELEMENT IS231C [Bacillus thuringiensis]
1253	Bt1Gc1596	Bt1G2555	1-757	g2226256	630	633	6.40E-62	51	51	(Y14084) hypothetical protein [Bacillus subtilis]
1253	Bt1Gc1596	Bt1G2556	759-1841	g2226255	1014	984	4.10E-99	56	100	(Y14084) hypothetical protein [Bacillus subtilis]
1253	Bt1Gc1596	Bt1G2557	2566-3973	g1177029	96	188	4.30E-14	41	20	HYPOTHETICAL 45.4 KD PROTEIN IN THIAMINASE I 5'REGION I]
1253	Bt1Gc1596	Bt1G2558	4232-3039	g3258413	305	403	1.50E-37	29	100	(AP000007) 404aa long hypothetical alany]-tRNA synthetase [Pyrococcus horikoshii]
1253	Bt1Gc1596	Bt1G2559	4846-6075	g1731082	942	931	1.70E-93	46	100	HYPOTHETICAL 44.7 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1253	Bt1Gc1596	Bt1G2560	6784-6638	g1621448	123	134	4.80E-09	56	98	(U72073) CotL [Bacillus subtilis]
1253	Bt1Gc1596	Bt1G2561	7288-7900	g2507080	275	342	4.40E-31	33	100	CHITOLIGOSACCHARIDE DEACETYLASE (NODULATION PROTEIN B) [Rhizobium sp. NGR234]
1253	Bt1Gc1596	Bt1G2562	8381-9274	g2506131	1207	1152	6.40E-117	79	33	ACONITATE HYDRATASE (CITRATE HYDRO-LYASE) (ACONITASE) [Bacillus subtilis]
1254	Bt1Gc1600	Bt1G2563	1362-1	g134740	147	216	9.80E-18	40	60	STAGE 0 SPORULATION PROTEIN B [Bacillus subtilis]
1254	Bt1Gc1600	Bt1G2564	671-390	g132830	433	433	9.90E-41	88	99	50S RIBOSOMAL PROTEIN L27 (BL30) (BL24) [Bacillus subtilis]
1254	Bt1Gc1600	Bt1G2565	1019-684	g141394	283	286	3.80E-25	49	99	HYPOTHETICAL 12.3 KD PROTEIN IN RPLU-RPMA INTERGENIC REGION (ORF X) [Bacillus subtilis]
1254	Bt1Gc1600	Bt1G2566	1339-1034	g132771	422	342	4.40E-31	80	99	50S RIBOSOMAL PROTEIN L21 (BL20) [Bacillus subtilis]
1254	Bt1Gc1600	Bt1G2567	2918-1250	g606187	576	528	6.30E-64	34	100	(U18997) ORF f495; orf of ECMRED, uses 2nd start [Escherichia coli]
1254	Bt1Gc1600	Bt1G2568	3827-2964	g134771	744	729	4.30E-72	48	100	STAGE IV SPORULATION PROTEIN FB [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1254	Bt1Gc1600	Bt1G2569	4425-3823	g134770	387	423	1.10E-39	48	77	STAGE IV SPORULATION PROTEIN FA [Bacillus subtilis]
1254	Bt1Gc1600	Bt1G2570	5497-4700	g400260	1015	875	1.40E-87	71	100	SEPTUM SITE-DETERMINING PROTEIN MIND [Bacillus subtilis]
1254	Bt1Gc1600	Bt1G2571	6186-5506	g400259	650	668	1.20E-65	60	100	SEPTUM SITE-DETERMINING PROTEIN MINC [Bacillus subtilis]
1254	Bt1Gc1600	Bt1G2572	6755-6240	g266570	307	182	3.90E-14	35	99	ROD SHAPE-DETERMINING PROTEIN MRED [Bacillus subtilis]
1254	Bt1Gc1600	Bt1G2573	7621-6773	g266569	623	606	4.60E-59	46	100	ROD SHAPE-DETERMINING PROTEIN MREC [Bacillus stearothermophilus]
1254	Bt1Gc1600	Bt1G2574	7989-7678	g266568	407	307	2.20E-27	79	31	ROD SHAPE-DETERMINING PROTEIN MREB [Bacillus subtilis]
1255	Bt1Gc1597	Bt1G2575	349-927	g1405459	588	554	1.50E-53	61	99	(Z73234) YneS [Bacillus subtilis]
1255	Bt1Gc1597	Bt1G2576	1051-1335	g1405458	277	277	3.40E-24	51	99	(Z73234) YneR [Bacillus subtilis]
1255	Bt1Gc1597	Bt1G2577	1736-1374	g1405456	406	406	7.20E-38	55	99	(Z73234) YneP [Bacillus subtilis]
1255	Bt1Gc1597	Bt1G2578	2325-1960	g1149666	433	433	9.90E-41	64	99	(X86498) IS1136 DNA [Clostridium perfringens]
1255	Bt1Gc1597	Bt1G2579	3078-873	g1787702	322	275	8.40E-37	31	98	(AE000240) putative virulence protein [Escherichia coli]
1255	Bt1Gc1597	Bt1G2580	3406-4686	g586861	1120	1069	4.00E-108	49	100	HYPOTHETICAL 48.6 KD PROTEIN IN SERS-DNAZ INTERGENIC REGION [Bacillus subtilis]
1255	Bt1Gc1597	Bt1G2581	6392-2362	g80261	221	292	2.70E-24	37	13	DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain (version 1) - Bacillus subtilis []
1255	Bt1Gc1597	Bt1G2582	7117-6539	g2634185	346	422	1.50E-39	45	99	(Z99113) similar to thiol:disulfide interchange protein [Bacillus subtilis]
1255	Bt1Gc1597	Bt1G2583	10276-7909	g1652519	317	350	2.50E-31	25	83	(D90906) ABC transporter [Synechocystis sp.]
1255	Bt1Gc1597	Bt1G2584	10276-9732	g2506131	756	771	1.50E-76	80	20	ACONITATE HYDRATASE (CITRATE HYDRO-LYASE) (ACONITASE) [Bacillus subtilis]
1256	Bt1Gc1602	Bt1G2585	811-257	g1706593	721	721	3.00E-71	74	99	ELONGATION FACTOR P (EF-P) [Bacillus subtilis]
1256	Bt1Gc1602	Bt1G2586	1891-836	g1731048	1223	1130	1.40E-114	65	100	PUTATIVE PEPTIDASE IN GCVT-SPOIIIAA INTERGENIC REGION [Bacillus subtilis]
1256	Bt1Gc1602	Bt1G2587	2339-1894	g1703004	443	392	2.20E-36	59	99	PUTATIVE CATABOLIC 3-DEHYDROQUINATE DEHYDRATASE (3-DEHYDROQUINASE) [Bacillus subtilis]
1256	Bt1Gc1602	Bt1G2588	3399-2398	g1731047	269	245	8.30E-21	34	99	HYPOTHETICAL 20.2 KD PROTEIN IN GCVT-SPOIIIAA INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1256	Bt1Gc1602	Bt1G2589	3122-4073	g1731046	1087	1085	8.10E-110	66	100	HYPOTHETICAL 36.0 KD PROTEIN IN GCVT-SPOIIIAA INTERGENIC REGION [Bacillus subtilis]
1256	Bt1Gc1602	Bt1G2590	4125-4523	g1731045	133	135	3.80E-09	32	99	HYPOTHETICAL 14.6 KD PROTEIN IN GCVT-SPOIIIAA INTERGENIC REGION [Bacillus subtilis]
1256	Bt1Gc1602	Bt1G2591	5188-4532	g1724017	448	457	2.80E-43	44	100	HYPOTHETICAL 24.7 KD PROTEIN IN CSPB-GLPP INTERGENIC REGION [Bacillus subtilis]
1256	Bt1Gc1602	Bt1G2592	5668-5243	g1731043	599	599	2.50E-58	79	99	HYPOTHETICAL 16.7 KD PROTEIN IN GCVT-SPOIIIAA INTERGENIC REGION [Bacillus subtilis]
1256	Bt1Gc1602	Bt1G2593	7536-5816	g586030	104	137	1.10E-08	40	38	SPORE PHOTOPRODUCT LYASE [Bacillus subtilis]
1256	Bt1Gc1602	Bt1G2594	6232-6574	g3582220	246	266	4.90E-23	43	99	(AE001272) conserved hypothetical protein [Lactococcus lactis]
1257	Bt1Gc1603	Bt1G2595	137-721	g1945679	350	375	1.40E-34	39	99	(Z94043) hypothetical protein [Bacillus subtilis]
1257	Bt1Gc1603	Bt1G2596	942-2141	g728970	1488	1194	2.30E-121	72	100	MULTIDRUG RESISTANCE PROTEIN 2 (MULTIDRUG-EFFLUX TRANSPORTER 2) [Bacillus subtilis]
1257	Bt1Gc1603	Bt1G2597	2514-2187	g1622733	200	233	1.60E-19	43	99	(U44893) orf108; unknown function [Butyrivibrio fibrisolvens]
1257	Bt1Gc1603	Bt1G2598	3271-3816	g730008	325	363	2.60E-33	38	99	MODULATOR OF DRUG ACTIVITY B [Escherichia coli]
1257	Bt1Gc1603	Bt1G2599	5620-3894	g2226188	524	684	2.50E-67	29	100	(Y14081) hypothetical protein [Bacillus subtilis]
1257	Bt1Gc1603	Bt1G2600	5264-3281	g2648784	131	248	2.20E-20	29	39	(AE000981) dipeptide ABC transporter, dipeptide-binding protein (dppA) [Archaeoglobus fulgidus]
1257	Bt1Gc1603	Bt1G2601	5722-6966	g2126617	522	508	1.10E-48	31	100	integral membrane protein LmrP - Lactococcus lactis [Lactococcus lactis]
1257	Bt1Gc1603	Bt1G2602	8066-7551	g2626825	609	615	5.10E-60	68	99	(D83967) YfkM [Bacillus subtilis]
1257	Bt1Gc1603	Bt1G2603	9134-8087	g1060858	773	820	9.70E-82	46	52	(D83026) KatE catalase [Bacillus subtilis]
1258	Bt1Gc1599	Bt1G2604	791-1	g1881327	862	742	1.80E-73	67	62	(AB001488) SIMILAR TO YDDS_BACSU. [Bacillus subtilis]
1258	Bt1Gc1599	Bt1G2606	3302-2649	g1817539	163	227	6.70E-19	28	99	(D78508) YfiO [Bacillus subtilis]
1258	Bt1Gc1599	Bt1G2607	6023-3521	g1652657	645	661	6.90E-65	46	57	(D90907) hypothetical protein [Synechocystis sp.]
1258	Bt1Gc1599	Bt1G2608	7761-5353	g4678633	202	279	7.50E-24	24	79	(AL049661) putative integral membrane efflux protein [Streptomyces coelicolor]
1258	Bt1Gc1599	Bt1G2609	6737-7202	g1945649	119	186	1.50E-14	30	94	(Z94043) hypothetical protein [Bacillus subtilis]
1258	Bt1Gc1599	Bt1G2610	8861-7224	g2634164	1184	1055	1.20E-106	43	100	(Z99113) yndJ [Bacillus subtilis]
1258	Bt1Gc1599	Bt1G2611	9821-8858	g2634163	263	326	2.20E-29	39	75	(Z99113) yndH [Bacillus subtilis]
1259	Bt1Gc1607	Bt1G2612	1022-1	g80410	214	269	2.40E-23	25	88	spore germination protein II - Bacillus subtilis (strain 168) []
1259	Bt1Gc1607	Bt1G2613	2579-982	g1805440	872	620	1.50E-60	37	97	(D50453) spore germination protein GerKA [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1259	Bt1Gc1607	Bt1G2614	2849-5786	g2635168	4439	4352	0	83	100	subtilis] similar to formate dehydrogenase [Bacillus subtilis]
1259	Bt1Gc1607	Bt1G2615	5805-6284	g3915537	491	441	1.40E-41	57	99	HYPOTHETICAL 17.5 KD PROTEIN IN SIGV-GREA INTERGENIC REGION [Bacillus subtilis]
1259	Bt1Gc1607	Bt1G2616	6403-6768	g3915538	428	434	7.80E-41	66	99	HYPOTHETICAL 14.0 KD PROTEIN IN SIGV-GREA INTERGENIC REGION [Bacillus subtilis]
1259	Bt1Gc1607	Bt1G2617	6761-7552	g730100	731	755	7.50E-75	55	100	NARQ PROTEIN [Bacillus subtilis]
1259	Bt1Gc1607	Bt1G2618	7569-8191	g730103	646	659	1.10E-64	60	61	NARA PROTEIN [Bacillus subtilis]
1260	Bt1Gc1604	Bt1G2619	964-317	g1945113	836	806	3.00E-80	76	100	(D88802) ydH [Bacillus subtilis]
1260	Bt1Gc1604	Bt1G2620	1146-3077	g3025120	2151	1937	2.10E-208	66	100	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YDIF [Bacillus subtilis]
1260	Bt1Gc1604	Bt1G2621	4579-3547	g3025119	1337	1299	1.70E-132	75	100	HYPOTHETICAL 36.8 KD PROTEIN IN PHOB-GROES INTERGENIC REGION [Bacillus subtilis]
1260	Bt1Gc1604	Bt1G2622	5022-4570	g1945109	442	449	2.00E-42	52	99	(D88802) H. influenzae, ribosomal protein alanine acetyltransferase; P44305 (189) [Bacillus subtilis]
1260	Bt1Gc1604	Bt1G2623	5722-5036	g1945108	738	738	4.80E-73	61	100	(D88802) H. influenzae hypothetical protein; P43990 (182) [Bacillus subtilis]
1260	Bt1Gc1604	Bt1G2624	6167-5769	g3323187	282	292	8.70E-26	43	99	(AE001257) conserved hypothetical protein [Treponema pallidum]
1261	Bt1Gc1609	Bt1G2625	766-4931	g2623773	966	1152	9.50E-115	37	11	(AF004835) tyrocidine synthetase 3 [Brevibacillus brevis]
1261	Bt1Gc1609	Bt1G2626	4989-5863	g585251	260	380	4.10E-35	31	58	POLYKETIDE SYNTHASE HETM [Anabaena sp.]
1262	Bt1Gc1606	Bt1G2627	3025-4329	g133292	835	855	1.90E-85	42	100	RNA POLYMERASE SIGMA-54 FACTOR [Bacillus subtilis]
1262	Bt1Gc1606	Bt1G2628	4723-5745	g465642	1113	950	1.60E-95	65	100	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN GAP 5'REGION (ORF1) [Bacillus megaterium]
1262	Bt1Gc1606	Bt1G2629	5775-6779	g421459	1386	1126	3.60E-114	81	100	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - Bacillus megaterium [Bacillus megaterium]
1262	Bt1Gc1606	Bt1G2630	6922-8903	g2120367	1748	1801	1.10E-187	55	100	phosphoglycerate kinase (EC 2.7.2.3) / triose-phosphate isomerase (EC 5.3.1.1) - Thermotoga maritima []
1262	Bt1Gc1606	Bt1G2631	8888-9454	g4589058	751	778	2.70E-77	77	37	(AF120091) 2,3-bisphosphoglycerate-independent phosphoglycerate mutase [Bacillus stearothermophilus]
1263	Bt1Gc1608	Bt1G2632	2196-1	g1495275	3744	3612	0	100	69	(X89230) DNA-directed RNA polymerase [Bacillus anthracis]
1263	Bt1Gc1608	Bt1G2633	3826-2191	g585920	2484	2397	7.50E-249	88	46	DNA-DIRECTED RNA POLYMERASE BETA CHAIN



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
										(TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Bacillus subtilis]
1264	Bt1Gc1605	Bt1G2634	474-1235	g1644216	725	717	8.00E-71	53	100	(D64126) unknown [Bacillus subtilis]
1264	Bt1Gc1605	Bt1G2635	1967-1374	g1708617	557	557	7.20E-54	52	99	KINB SIGNALING PATHWAY ACTIVATION PROTEIN [Bacillus subtilis]
1264	Bt1Gc1605	Bt1G2636	2107-2664	g121132	424	435	6.10E-41	42	99	SPORE GERMINATION PROTEIN GERD PRECURSOR [Bacillus subtilis]
1264	Bt1Gc1605	Bt1G2637	3963-2902	g1723296	1238	1121	1.20E-113	66	100	HYPOTHETICAL 38.6 KD PROTEIN IN CWLD-GERD INTERGENIC REGION [Bacillus subtilis]
1264	Bt1Gc1605	Bt1G2638	4824-4111	g1706210	727	619	1.90E-60	59	100	GERMINATION-SPECIFIC N-ACETYLMURAMOYL-L-ALANINE AMIDASE (CELL WALL HYDROLASE) (AUTOLYSIN) [Bacillus subtilis]
1264	Bt1Gc1605	Bt1G2639	5316-4891	g1723295	204	189	7.10E-15	29	99	HYPOTHETICAL 17.6 KD PROTEIN IN CWLD 5'REGION (ORF1) [Bacillus subtilis]
1264	Bt1Gc1605	Bt1G2640	5871-5482	g2507325	548	548	6.50E-53	82	99	30S RIBOSOMAL PROTEIN S9 (BS10) [Bacillus subtilis]
1264	Bt1Gc1605	Bt1G2641	6330-5896	g2500252	621	621	1.20E-60	82	99	50S RIBOSOMAL PROTEIN L13 [Bacillus subtilis]
1264	Bt1Gc1605	Bt1G2642	7223-6486	g3915172	718	703	2.40E-69	56	100	TRNA PSEUDOURIDINE SYNTHASE A (PSEUDOURIDYLATE SYNTHASE I) (URACIL HYDROLYASE) [Bacillus subtilis]
1264	Bt1Gc1605	Bt1G2643	7814-7244	g2632414	669	544	1.70E-52	68	72	(Z99104) ybaF [Bacillus subtilis]
1265	Bt1Gc1610	Bt1G2644	632-54	g2415741	708	652	6.20E-64	68	99	(AB000617) YceD [Bacillus subtilis]
1265	Bt1Gc1610	Bt1G2645	1256-663	g2415740	649	582	1.60E-56	60	99	(AB000617) YceC [Bacillus subtilis]
1265	Bt1Gc1610	Bt1G2646	1924-1385	g141085	292	283	7.80E-25	36	99	HYPOTHETICAL PROTEIN IN NAGH 5'REGION (ORFA) [Clostridium perfringens]
1265	Bt1Gc1610	Bt1G2648	4938-3523	g2688299	1163	1127	2.90E-114	48	100	(AE001145) prolyl-rRNA synthetase (proS) [Borrelia burgdorferi]
1265	Bt1Gc1610	Bt1G2649	5691-5307	g2633006	116	107	3.50E-06	22	99	(Z99107) yesK [Bacillus subtilis]
1265	Bt1Gc1610	Bt1G2650	4116-7623	g219944	101	246	1.20E-19	27	23	(D12485) The first in-frame ATG codon is located at nucleotides 17-19, followed by a second ATG codon 52 codons downstream. And the second ATG codon is potential initiation point for translation of NPPase. [Homo sapiens]
1265	Bt1Gc1610	Bt1G2651	7165-7623	g1177011	501	371	3.70E-34	58	36	HYPOTHETICAL 47.3 KD PROTEIN IN WAPA-LICT INTERGENIC REGION [Bacillus subtilis]
1266	Bt1Gc1611	Bt1G2652	2231-768	g1730257	2232	2137	2.70E-221	86	100	PROBABLE GLYCINE DEHYDROGENASE

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1266	Bt1Gc1611	Bt1G2653	3571-2227	g1730256	1689	1600	2.20E-164	73	100	(DECARBOXYLATING) SUBUNIT 2 (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-PROTEIN) [Bacillus subtilis]
1266	Bt1Gc1611	Bt1G2654	4686-3601	g1730258	1274	1243	1.50E-126	67	100	PROBABLE GLYCINE DEHYDROGENASE (DECARBOXYLATING) SUBUNIT 1 (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-PROTEIN) [Bacillus subtilis]
1266	Bt1Gc1611	Bt1G2655	5083-6753	g1731040	1976	1996	2.30E-206	67	100	PROBABLE AMINOMETHYLTRANSFERASE (GLYCINE CLEAVAGE SYSTEM T PROTEIN) [Bacillus subtilis]
1266	Bt1Gc1611	Bt1G2656	6752-7236	g1731039	449	475	3.50E-45	51	61	HYPOTHETICAL HELICASE IN SINI-GCVT INTERGENIC REGION [Bacillus subtilis]
1267	Bt1Gc1612	Bt1G2657	295-1089	g3025279	873	716	1.00E-70	63	100	HYPOTHETICAL 30.8 KD PROTEIN IN SINI-GCVT INTERGENIC REGION [Bacillus subtilis]
1267	Bt1Gc1612	Bt1G2658	1099-2115	g2633798	989	950	1.60E-95	55	100	HYPOTHETICAL 28.5 KD PROTEIN IN SIGV-GREA INTERGENIC REGION [Bacillus subtilis]
1267	Bt1Gc1612	Bt1G2659	2130-3419	g2633799	1359	1348	1.10E-137	61	100	(Z99111) molybdopterin biosynthesis protein [Bacillus subtilis]
1267	Bt1Gc1612	Bt1G2660	3427-3897	g2633801	532	506	1.80E-48	67	99	(Z99111) molybdopterin biosynthesis protein [Bacillus subtilis]
1267	Bt1Gc1612	Bt1G2661	3890-4123	g2633802	167	185	1.90E-14	48	99	(Z99111) molybdopterin converting factor (subunit 1) [Bacillus subtilis]
1267	Bt1Gc1612	Bt1G2662	4199-5259	g1723296	715	796	3.40E-79	44	100	HYPOTHETICAL 38.6 KD PROTEIN IN CWLD-GERD INTERGENIC REGION [Bacillus subtilis]
1267	Bt1Gc1612	Bt1G2663	6110-6365	g1770010	181	169	9.40E-13	43	99	(Z75208) hypothetical protein [Bacillus subtilis]
1267	Bt1Gc1612	Bt1G2664	6460-7257	g2634341	675	668	1.20E-65	53	100	(Z99114) similar to hypothetical proteins [Bacillus subtilis]
1267	Bt1Gc1612	Bt1G2665	7466-7906	g1652288	262	209	5.40E-17	40	99	(D90904) hypothetical protein [Synechocystis sp.]
1267	Bt1Gc1612	Bt1G2666	7997-9747	g2619017	1036	1172	4.90E-119	43	100	(AF027868) squalene-hopene cyclase [Bacillus subtilis]
1268	Bt1Gc1613	Bt1G2667	7548-7934	g2415388	128	123	7.00E-08	46	99	(AF015775) YodI [Bacillus subtilis]
1268	Bt1Gc1613	Bt1G2668	8691-7981	g2415387	317	330	8.20E-30	32	100	(AF015775) YodH [Bacillus subtilis]
1268	Bt1Gc1613	Bt1G2669	10465-8792	g2632017	811	846	1.70E-84	35	99	(AJ002571) DppE [Bacillus subtilis]
1269	Bt1Gc1615	Bt1G2670	1379-966	g120294	226	278	2.60E-24	43	99	FLAGELLAR BASAL-BODY ROD PROTEIN FLGC [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1269	Bt1Gc1615	Bt1G2671	1814-1398	g1272365	82	163	4.10E-12	28	99	(U51896) LfgB [Vibrio parahaemolyticus]
1269	Bt1Gc1615	Bt1G2672	5083-4221	g2829830	236	299	1.60E-26	27	100	FLAGELLAR HOOK-ASSOCIATED PROTEIN 3 (HAP3) [Bacillus subtilis]
1269	Bt1Gc1615	Bt1G2673	6439-4138	g2314271	120	208	6.70E-16	24	36	(AE000618) flagellar hook-associated protein 1 (HAP1) (flgK) [Helicobacter pylori 26695]
1269	Bt1Gc1615	Bt1G2674	6436-3613	g3322960	126	269	1.90E-22	23	54	(AE001240) flagellar hook-associated protein 1 (flgK) [Treponema pallidum]
1269	Bt1Gc1615	Bt1G2675	8074-8850	g143805	595	616	4.00E-60	47	100	(M80245) CheR [Bacillus subtilis]
1269	Bt1Gc1615	Bt1G2676	11350-10272	g2983912	113	91	0.00017	33	99	(AE000745) flagellar switch protein FlitN [Aquifex aeolicus]
1269	Bt1Gc1615	Bt1G2677	12502-11939	g1556378	833	780	1.70E-77	89	99	(Y08031) cheD [Bacillus cereus]
1269	Bt1Gc1615	Bt1G2678	12502-11939	g1556378	833	780	1.70E-77	89	99	(Y08031) cheD [Bacillus cereus]
1269	Bt1Gc1615	Bt1G2679	13200-12837	g2500757	258	284	4.30E-24	54	19	CHEMOTAXIS PROTEIN CHEA [Listeria monocytogenes]
1270	Bt1Gc1614	Bt1G2680	1-805	g1619836	963	881	3.30E-88	79	71	(U67921) REP [Bacillus thuringiensis israelensis]
1270	Bt1Gc1614	Bt1G2681	4123-5322	g80336	439	428	3.40E-40	32	100	Mob protein - Bacillus sp. plasmid pTB19 [Plasmid pTB19]
1270	Bt1Gc1614	Bt1G2682	7349-8116	g1619836	553	393	1.70E-36	67	53	(U67921) REP [Bacillus thuringiensis israelensis]
1271	Bt1Gc1617	Bt1G2683	267-1151	g4160473	306	384	1.50E-35	37	100	(AF109909) PhaR [Bacillus megaterium]
1271	Bt1Gc1617	Bt1G2684	912-1652	g4160474	911	782	1.00E-77	73	100	(AF109909) 3-ketoacyl-CoA reductase PhaB [Bacillus megaterium]
1271	Bt1Gc1617	Bt1G2685	3295-3897	g2492902	385	391	2.80E-36	39	100	HYPOTHETICAL 21.7 KD PROTEIN IN BETT-PRPR INTERGENIC REGION [Escherichia coli]
1271	Bt1Gc1617	Bt1G2686	4756-3920	g1763705	433	473	5.70E-45	37	100	(Z83337) ywpD [Bacillus subtilis]
1271	Bt1Gc1617	Bt1G2687	5741-5009	g1652287	269	330	1.30E-28	31	27	(D90904) sensory transduction histidine kinase [Synecocystis sp.]
1272	Bt1Gc1618	Bt1G2688	1127-219	g4056456	247	400	3.10E-37	32	100	(AC005990) Strong similarity to gb U20808 auxin-induced protein from Vigna radiata and a member of the zinc-binding dehydrogenase family PF00107. ESTs gb T43674, gb H77006 and gb AA395179 come from this gene. [Arabidopsis thaliana]
1272	Bt1Gc1618	Bt1G2689	1127-1	g3080402	204	392	2.20E-36	33	92	(AL022603) putative NADPH quinone oxidoreductase [Arabidopsis thaliana]
1272	Bt1Gc1618	Bt1G2690	2140-3546	g1620930	1363	1389	4.90E-142	58	100	(Z79580) putative orf [Bacillus subtilis]
1272	Bt1Gc1618	Bt1G2691	3585-4076	g1620931	401	405	9.20E-38	42	99	(Z79580) putative orf [Bacillus subtilis]
1272	Bt1Gc1618	Bt1G2692	6024-7226	g1881374	892	804	4.80E-80	46	100	(AB001488) SIMILAR TO BICYCLOMYCIN

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1272	Bt1Gc1618	Bt1G2693	8717-9610	g2632518	778	748	4.10E-74	51	100	RESISTANCE PROTEIN. [Bacillus subtilis] (Z99105) similar to transcriptional regulator (AraC/XylIS family) [Bacillus subtilis]
1273	Bt1Gc1619	Bt1G2694	1-367	g3005554	100	138	7.20E-09	24	37	(AF047044) putative transposase [Anabaena PCC7120]
1273	Bt1Gc1619	Bt1G2695	7479-7201	g39805	218	252	1.50E-21	53	99	(X61953) abrB [Bacillus subtilis]
1274	Bt1Gc1620	Bt1G2696	12-950	g2633168	280	398	5.10E-37	32	100	(Z99108) similar to iron(III) dicitrate transport permease [Bacillus subtilis]
1274	Bt1Gc1620	Bt1G2697	1503-1000	g1941918	240	283	7.80E-25	35	99	(X93081) sigma F/sigma G transcribed gene [Bacillus subtilis]
1274	Bt1Gc1620	Bt1G2698	6453-5350	g2635250	1329	1267	4.20E-129	67	100	(Z99118) quinolinate synthetase [Bacillus subtilis]
1274	Bt1Gc1620	Bt1G2699	7317-6448	g3183539	843	794	5.50E-79	57	100	PROBABLE NICOTINATE-NUCLEOTIDE PYROPHOSPHORYLASE (CARBOXYLATING) (QUINOLINATE PHOSPHORIBOSYLTRANSFERASE (DECARBOXYLATING)) (QAPRTASE) (GENERAL STRESS PROTEIN 70) (GSP70) [Bacillus subtilis]
1274	Bt1Gc1620	Bt1G2700	8866-7021	g585537	1051	1129	1.80E-114	43	100	L-ASPARTATE OXIDASE (QUINOLINATE SYNTHETASE B) [Bacillus subtilis]
1274	Bt1Gc1620	Bt1G2701	9022-9516	g585557	388	400	3.10E-37	50	42	NIFS PROTEIN HOMOLOG [Bacillus subtilis]
1275	Bt1Gc1621	Bt1G2702	123-2195	g1865711	600	741	2.30E-73	47	87	(Y11477) endolysin [Bacteriophage Bastille]
1275	Bt1Gc1621	Bt1G2703	2499-3260	g2612912	828	519	7.70E-50	63	100	(AF015825) unknown [Bacillus subtilis]
1275	Bt1Gc1621	Bt1G2704	4568-4197	g481913	156	147	2.00E-10	34	99	hypothetical protein - Clostridium butyricum [Clostridium butyricum]
1275	Bt1Gc1621	Bt1G2705	6597-4642	g135956	459	718	6.30E-71	29	100	TRANSPOSASE B (TRANSPOSON TN554) [Staphylococcus aureus]
1275	Bt1Gc1621	Bt1G2706	6610-4600	g481912	478	708	7.20E-70	28	100	hypothetical protein - Clostridium butyricum [Clostridium butyricum]
1275	Bt1Gc1621	Bt1G2707	7670-6579	g224807	555	648	1.60E-63	37	100	gene tnpA [Staphylococcus aureus]
1275	Bt1Gc1621	Bt1G2708	8057-9110	g4512351	275	341	5.60E-31	24	98	(AB011836) response regulator aspartate phosphatase [Bacillus halodurans]
1276	Bt1Gc1624	Bt1G2709	1075-1	g1256135	854	681	5.20E-67	51	80	(AB002150) YbbF [Bacillus subtilis]
1276	Bt1Gc1624	Bt1G2710	1990-1079	g2632437	712	613	8.40E-60	49	100	(Z99104) similar to hypothetical proteins [Bacillus subtilis]
1276	Bt1Gc1624	Bt1G2711	3140-2301	g141086	308	331	6.40E-30	27	100	HYPOTHETICAL 31.2 KD PROTEIN IN NAGH 5'REGION (ORFB) [Clostridium perfringens]
1276	Bt1Gc1624	Bt1G2712	3152-2292	g2507342	244	289	1.80E-25	27	100	HYPOTHETICAL PROTEIN HI0143 [Haemophilus influenzae Rd]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1276	Bt1Gc1624	Bt1G2713	4745-3591	g1098508	1159	817	2.00E-81	58	100	(U17283) putative spore germination apparatus protein [Bacillus megaterium]
1276	Bt1Gc1624	Bt1G2714	5001-6371	g2633047	1445	1352	4.10E-138	57	100	(Z99107) similar to metabolite transporter [Bacillus subtilis]
1276	Bt1Gc1624	Bt1G2715	11624-10944	g2633697	604	556	9.20E-54	55	100	(Z99110) similar to alkaline phosphatase [Bacillus subtilis]
1277	Bt1Gc1616	Bt1G2716	1-230	g1169301	277	293	6.80E-26	70	29	2,3-DIHYDRO-2,3-DIHYDROXYBENZOATE DEHYDROGENASE (COLD SHOCK PROTEIN CSI14) [Bacillus subtilis]
1277	Bt1Gc1616	Bt1G2717	256-1440	g1169303	1121	1107	3.80E-112	59	100	ISOCHORISMATE SYNTHASE DHBC [Bacillus subtilis]
1277	Bt1Gc1616	Bt1G2718	1471-3087	g1169324	2036	1949	2.20E-201	70	100	2,3-DIHYDROXYBENZOATE-AMP LIGASE (DIHYDROXYBENZOIC ACID-ACTIVATING ENZYME) [Bacillus subtilis]
1277	Bt1Gc1616	Bt1G2719	4039-10381	g4481934	4029	2965	0	43	30	(AL035640) CDA peptide synthetase I [Streptomyces coelicolor]
1277	Bt1Gc1616	Bt1G2721	11193-11405	g4337034	206	218	6.00E-18	52	99	(AF124138) Cda-orfX [Streptomyces coelicolor]
1277	Bt1Gc1616	Bt1G2722	11413-11752	g401630	244	254	3.60E-21	41	24	HYPOTHETICAL 51.5 KD PROTEIN IN RBSR-RRSC INTERGENIC REGION [Escherichia coli]
1278	Bt1Gc1622	Bt1G2723	368-1990	g2635784	2066	1717	8.60E-177	73	100	(Z99120) similar to multidrug-efflux transporter [Bacillus subtilis]
1278	Bt1Gc1622	Bt1G2724	2128-3105	g2851638	686	716	1.00E-70	42	100	RIBOSE OPERON REPRESSOR [Bacillus subtilis]
1278	Bt1Gc1622	Bt1G2725	3113-3997	g3915836	814	705	1.50E-69	57	100	RIBOKINASE [Bacillus subtilis]
1278	Bt1Gc1622	Bt1G2726	4006-4398	g548707	388	388	5.80E-36	60	99	HIGH AFFINITY RIBOSE TRANSPORT PROTEIN RBSD [Bacillus subtilis]
1278	Bt1Gc1622	Bt1G2727	4420-5901	g1894759	1523	1541	3.80E-158	61	100	(Z92953) ATP-binding transport protein [Bacillus subtilis]
1278	Bt1Gc1622	Bt1G2728	5873-6757	g1894758	1014	628	1.60E-77	69	91	(Z92953) membrane transport protein [Bacillus subtilis]
1279	Bt1Gc1625	Bt1G2729	522-1	g1256141	169	204	1.80E-16	28	94	(AB002150) YbbL [Bacillus subtilis]
1279	Bt1Gc1625	Bt1G2730	2540-873	g2851477	2249	2147	2.30E-222	75	100	ARGINYL-TRNA SYNTHETASE (ARGININE--TRNA LIGASE) (ARGRS) [Bacillus subtilis]
1279	Bt1Gc1625	Bt1G2731	2968-2538	g2224755	196	244	1.10E-20	31	99	(Z97024) ywiB [Bacillus subtilis]
1279	Bt1Gc1625	Bt1G2732	3180-4163	g2619011	910	889	4.70E-89	55	100	(AF027868) YocD [Bacillus subtilis]
1279	Bt1Gc1625	Bt1G2733	4787-5815	g280310	1029	400	1.30E-61	58	100	regulatory protein pfoR - Clostridium perfringens []
1279	Bt1Gc1625	Bt1G2734	7044-6174	g1565242	1221	1228	5.70E-125	80	100	(Z80360) Unknown, highly similar to several agmatinases [Bacillus subtilis]
1279	Bt1Gc1625	Bt1G2735	8092-7265	g1565241	1259	1194	2.30E-121	84	100	(Z80360) Unknown, highly similar to several spermidine synthases [Bacillus subtilis]
1280	Bt1Gc1623	Bt1G2736	54-689	g1592698	770	724	1.40E-71	69	100	(Y08559) Unknown [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1280	Bt1Gc1623	Bt1G2737	6379-5540	g2293328	1137	1062	2.20E-107	76	100	(AF008220) putative morphine dehydrogenase [Bacillus subtilis]
1280	Bt1Gc1623	Bt1G2738	7580-6393	g2293329	1541	1270	2.00E-129	75	100	(AF008220) YtbD [Bacillus subtilis]
1280	Bt1Gc1623	Bt1G2739	7944-8318	g2293276	414	421	1.90E-39	70	99	(AF008220) YtcD [Bacillus subtilis]
1281	Bt1Gc1628	Bt1G2740	1-1158	g1934815	1385	1202	3.20E-122	67	86	(Z93937) unknown [Bacillus subtilis]
1281	Bt1Gc1628	Bt1G2741	1227-2824	g2293261	1234	1170	7.90E-119	62	100	(AF008220) YtsJ [Bacillus subtilis]
1281	Bt1Gc1628	Bt1G2742	4344-3419	g401589	445	494	3.40E-47	33	100	HYPOTHEICAL 33.1 KD PROTEIN IN SELC-NLPA INTERGENIC REGION [Escherichia coli]
1281	Bt1Gc1628	Bt1G2743	4471-5859	g1881332	922	937	3.90E-94	43	100	(AB001488) SIMILAR TO THE RHIZOPINE CATABOLISM (MOCR) GENE OF RHIZOBIUM MELIOTI. [Bacillus subtilis]
1281	Bt1Gc1628	Bt1G2744	8025-6562	g3687664	819	837	1.50E-83	38	100	(AF049873) sensor protein [Lactococcus lactis]
1281	Bt1Gc1628	Bt1G2745	8661-8076	g3687663	327	414	1.00E-38	42	85	(AF049873) response protein [Lactococcus lactis]
1282	Bt1Gc1629	Bt1G2746	454-1	g133449	360	285	4.80E-25	50	89	DNA-DIRECTED RNA POLYMERASE DELTA SUBUNIT [Bacillus subtilis]
1282	Bt1Gc1629	Bt1G2747	1220-1	g2983006	95	154	3.60E-11	30	61	(AE000683) putative protein [Aquifex aeolicus]
1282	Bt1Gc1629	Bt1G2748	2444-1308	g1168281	1587	1443	9.30E-148	80	100	ACYL-COA DEHYDROGENASE [Bacillus subtilis]
1282	Bt1Gc1629	Bt1G2749	3590-2457	g1168282	1275	1250	2.60E-127	66	100	ACYL-COA DEHYDROGENASE [Bacillus subtilis]
1282	Bt1Gc1629	Bt1G2750	4521-3661	g1170972	958	965	4.20E-97	64	100	PROBABLE 3-HYDROXYBUTYRYL-COA DEHYDROGENASE (BETA-HYDROXYBUTYRYL-COA DEHYDROGENASE) (BHBD) [Bacillus subtilis]
1282	Bt1Gc1629	Bt1G2751	5752-4574	g1303931	1252	1172	4.90E-119	61	100	(D84432) YqIL [Bacillus subtilis]
1282	Bt1Gc1629	Bt1G2752	7975-5858	g1176956	2906	2931	1.90E-305	77	100	HYPOTHEICAL 79.2 KD PROTEIN IN ACDA 5'REGION [Bacillus subtilis]
1283	Bt1Gc1627	Bt1G2753	3015-348	g2337795	1509	1630	1.40E-167	40	100	(Y13937) putative PacL protein [Bacillus subtilis]
1283	Bt1Gc1627	Bt1G2754	7495-6035	g98699	1565	1538	8.00E-158	60	100	catalase (EC 1.11.1.6) - Listeria seeligeri []
1283	Bt1Gc1627	Bt1G2755	8514-7591	g417116	1045	1057	7.50E-107	61	100	FERROCHELATA TASE (PROTOHEME FERRO-LYASE) (HEME SYNTHETASE) [Bacillus subtilis]
1284	Bt1Gc1631	Bt1G2756	494-59	g1770003	426	373	2.30E-34	62	99	(Z75208) hypothetical protein [Bacillus subtilis]
1284	Bt1Gc1631	Bt1G2757	1367-630	g1770002	797	827	1.80E-82	66	100	(Z75208) autolysin response regulator [Bacillus subtilis]
1284	Bt1Gc1631	Bt1G2758	3114-1348	g1770001	1944	1747	5.70E-180	65	100	(Z75208) autolysin sensor kinase [Bacillus subtilis]
1284	Bt1Gc1631	Bt1G2759	3466-4734	g1074652	899	809	1.40E-80	44	100	hypothetical protein H1104 - Haemophilus influenzae (strain Rd KW20) [Haemophilus influenzae Rd]
1284	Bt1Gc1631	Bt1G2760	6267-4331	g4835822	632	617	3.20E-60	33	95	(AF102174) glycine betaine transporter BetL [Listeria monocytogenes]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>1</sub> nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1285	Bt1Gc1630	Bt1G2761	79-1322	g732364	807	735	9.90E-73	43	91	HYPOTHETICAL 50.0 KD PROTEIN IN UNG-ROCA INTERGENIC REGION [Bacillus subtilis]
1285	Bt1Gc1630	Bt1G2762	1591-2394	g401595	456	412	1.70E-38	37	100	HYPOTHETICAL 29.7 KD PROTEIN IN IBPA-GYRB INTERGENIC REGION [Escherichia coli]
1285	Bt1Gc1630	Bt1G2763	3634-2960	g731058	770	770	1.90E-76	64	100	URACIL-DNA GLYCOSYLASE (UDG) [Bacillus subtilis]
1285	Bt1Gc1630	Bt1G2764	5556-4627	g1724002	701	743	1.40E-73	47	100	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN 2 IN GLPD-CSPB INTERGENIC REGION [Bacillus subtilis]
1285	Bt1Gc1630	Bt1G2765	6322-5568	g1730938	278	348	1.00E-31	31	100	HYPOTHETICAL 28.9 KD PROTEIN IN ILVA 3'REGION [Bacillus subtilis]
1285	Bt1Gc1630	Bt1G2766	7422-6448	g2633758	643	637	2.40E-62	40	100	(Z99111) similar to transcriptional regulator (LacI family) [Bacillus subtilis]
1285	Bt1Gc1630	Bt1G2767	9471-8177	g118592	1513	1452	1.00E-148	68	100	HOMOSERINE DEHYDROGENASE (HDH) [Bacillus subtilis]
1286	Bt1Gc1632	Bt1G2768	497-323	g2634358	151	162	5.20E-12	52	98	(Z99114) yozD [Bacillus subtilis]
1286	Bt1Gc1632	Bt1G2769	1466-816	g2415400	392	410	2.70E-38	38	100	(AF015775) YodN [Bacillus subtilis]
1286	Bt1Gc1632	Bt1G2770	1932-1714	g2634360	163	185	1.90E-14	49	99	(Z99114) yozE [Bacillus subtilis]
1286	Bt1Gc1632	Bt1G2771	2301-2008	g2529468	310	270	1.90E-23	60	99	(AF006665) Yoku [Bacillus subtilis]
1286	Bt1Gc1632	Bt1G2772	3698-2286	g4033499	2025	1949	2.20E-201	77	100	HYPOTHETICAL 54.1 KD PROTEIN IN DEOD-ARGE INTERGENIC REGION [Bacillus subtilis]
1286	Bt1Gc1632	Bt1G2773	5294-3909	g1064808	843	870	4.90E-87	42	100	(D78193) positive regulatory protein [Bacillus subtilis]
1286	Bt1Gc1632	Bt1G2774	6584-5522	g2415402	434	472	7.30E-45	39	100	(AF015775) YodP [Bacillus subtilis]
1286	Bt1Gc1632	Bt1G2775	7623-6322	g2529465	1371	1308	1.90E-133	56	100	(AF006665) YokuP [Bacillus subtilis]
1286	Bt1Gc1632	Bt1G2776	10045-7630	g4557817	757	683	5.80E-88	39	100	3-oxoacid CoA transferase precursor; succinyl-CoA:3-ketoacid-CoA transferase precursor [Homo sapiens]
1286	Bt1Gc1632	Bt1G2777	10291-8990	g2529462	1096	1076	7.20E-109	51	100	(AF006665) YokuM [Bacillus subtilis]
1286	Bt1Gc1632	Bt1G2778	13091-12027	g2127280	1831	1831	7.10E-189	100	95	histidine protein kinase (Tn5401) - Bacillus thuringiensis [Bacillus thuringiensis]
1287	Bt1Gc1633	Bt1G2779	659-99	g1731021	430	352	3.80E-32	45	99	HYPOTHETICAL 21.4 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
1287	Bt1Gc1633	Bt1G2780	3346-1570	g1653786	880	942	1.10E-94	32	71	(D90916) mannose-1-phosphate guanyltransferase [Synechocystis sp.]
1287	Bt1Gc1633	Bt1G2781	4334-3667	g2983430	326	286	3.80E-25	37	95	(AE000713) transcriptional regulator (PhoU-like) [Aquifex aeolicus]
1287	Bt1Gc1633	Bt1G2782	5393-4638	g3024474	887	890	3.70E-89	65	100	PROBABLE PHOSPHATE TRANSPORT ATP-BINDING

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1287	Bt1Gc1633	Bt1G2783	6309-5472	g2649217	548	585	7.80E-57	43	100	PROTEIN PSTB [Methanococcus jannaschii] (AE001010) phosphate ABC transporter, permease protein (pstA) [Archaeoglobus fulgidus]
1287	Bt1Gc1633	Bt1G2784	7222-6347	g2649218	655	685	2.00E-67	45	100	(AE001010) phosphate ABC transporter, permease protein (pstC) [Archaeoglobus fulgidus]
1287	Bt1Gc1633	Bt1G2785	8239-7262	g2498946	428	455	4.60E-43	30	100	SPHX PROTEIN HOMOLOG PRECURSOR [Synecocystis sp.]
1287	Bt1Gc1633	Bt1G2786	10639-8562	g1731017	1750	1718	6.70E-177	53	97	HYPOTHETICAL 80.1 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
1288	Bt1Gc1636	Bt1G2787	601-1209	g1929333	578	590	2.30E-57	54	100	(Z93767) ywrF [Bacillus subtilis]
1288	Bt1Gc1636	Bt1G2788	2695-2146	g732385	497	291	1.10E-25	55	99	HYPOTHETICAL 19.7 KD PROTEIN IN SPOIR-GLYC INTERGENIC REGION [Bacillus subtilis]
1288	Bt1Gc1636	Bt1G2789	3809-2775	g732384	1091	1103	1.00E-111	62	100	HYPOTHETICAL 37.0 KD PROTEIN IN SPOIR-GLYC INTERGENIC REGION [Bacillus subtilis]
1288	Bt1Gc1636	Bt1G2790	5289-4432	g730770	520	511	5.40E-49	52	100	STAGE II SPORULATION PROTEIN R [Bacillus subtilis]
1288	Bt1Gc1636	Bt1G2791	6256-5396	g1170229	758	687	1.20E-67	52	100	HEMK PROTEIN HOMOLOG [Bacillus subtilis]
1288	Bt1Gc1636	Bt1G2792	7320-6253	g1172909	1551	1459	1.90E-149	82	100	PEPTIDE CHAIN RELEASE FACTOR 1 (RF-1) [Bacillus subtilis]
1288	Bt1Gc1636	Bt1G2793	8174-7590	g729902	720	720	3.80E-71	69	99	THYMIDINE KINASE [Bacillus subtilis]
1288	Bt1Gc1636	Bt1G2794	8528-8286	g4193373	321	321	7.30E-29	73	99	(AF072894) ribosomal protein L31 [Listeria monocytogenes]
1289	Bt1Gc1626	Bt1G2795	273-1	g2293166	337	342	4.40E-31	73	19	(AF008220) amino acid transporter [Bacillus subtilis]
1289	Bt1Gc1626	Bt1G2796	560-1852	g2293313	1325	1118	2.60E-113	58	100	(AF008220) YtiP [Bacillus subtilis]
1289	Bt1Gc1626	Bt1G2797	5312-3486	g2633518	1537	1398	5.50E-143	53	100	(Z99110) similar to Na <sup>+</sup> /H <sup>+</sup> antiporter [Bacillus subtilis]
1289	Bt1Gc1626	Bt1G2798	8404-6461	g464512	2555	2265	7.30E-235	72	100	QUINOL OXIDASE POLYPEPTIDE I (QUINOL OXIDASE AA3-600, SUBUNIT QOXB) (OXIDASE AA(3) SUBUNIT 1) [Bacillus subtilis]
1289	Bt1Gc1626	Bt1G2799	9310-8357	g2636352	705	692	3.60E-68	47	100	(Z99123) cytochrome aa3 quinol oxidase (subunit II) [Bacillus subtilis]
1289	Bt1Gc1626	Bt1G2800	10999-9602	g730600	1542	1403	1.60E-143	61	100	AMINO-ACID PERMEASE ROCC [Bacillus subtilis]
1290	Bt1Gc1637	Bt1G2801	1-227	g39805	310	310	1.10E-27	79	80	(X61953) abrB [Bacillus subtilis]
1290	Bt1Gc1637	Bt1G2802	1129-257	g586874	914	868	8.00E-87	62	100	HYPOTHETICAL 33.0 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION [Bacillus subtilis]
1290	Bt1Gc1637	Bt1G2803	2115-1375	g586873	977	978	1.80E-98	73	100	HYPOTHETICAL 28.3 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION [Bacillus subtilis]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1290	Bt1Gc1637	Bt1G2804	2545-2198	g586872	378	411	2.10E-38	60	99	HYPOTHETICAL 14.1 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION [Bacillus subtilis]
1290	Bt1Gc1637	Bt1G2805	3387-2563	g586871	1147	1147	2.20E-116	76	100	HYPOTHETICAL 31.2 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION [Bacillus subtilis]
1290	Bt1Gc1637	Bt1G2806	4376-3396	g586870	862	816	2.60E-81	53	100	DNA POLYMERASE III, DELTA' SUBUNIT [Bacillus subtilis]
1290	Bt1Gc1637	Bt1G2807	5038-4403	g586867	579	591	1.80E-57	52	100	THYMIDYLATE KINASE (DTMP KINASE) [Bacillus subtilis]
1290	Bt1Gc1637	Bt1G2808	6452-4943	g586866	1006	1079	3.50E-109	49	100	HYPOTHETICAL 53.2 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION [Bacillus subtilis]
1291	Bt1Gc1638	Bt1G2809	1-863	g586901	1183	1201	4.10E-122	81	63	DNA REPAIR PROTEIN RADA HOMOLOG (DNA REPAIR PROTEIN SMS HOMOLOG) [Bacillus subtilis]
1291	Bt1Gc1638	Bt1G2810	870-1940	g586902	1136	1087	4.90E-110	63	100	HYPOTHETICAL 40.7 KD PROTEIN IN MECB-GLTX INTERGENIC REGION [Bacillus subtilis]
1291	Bt1Gc1638	Bt1G2811	2104-3210	g586903	1252	1243	1.50E-126	70	100	HYPOTHETICAL 40.9 KD PROTEIN IN MECB-GLTX INTERGENIC REGION [Bacillus subtilis]
1291	Bt1Gc1638	Bt1G2812	3227-3919	g2127058	579	603	9.60E-59	51	100	hypothetical protein - Bacillus subtilis [Bacillus subtilis]
1291	Bt1Gc1638	Bt1G2813	4027-4500	g549577	645	647	2.10E-63	76	99	HYPOTHETICAL 17.1 KD PROTEIN IN MECB-GLTX INTERGENIC REGION [Bacillus subtilis]
1291	Bt1Gc1638	Bt1G2814	4593-6041	g135106	1927	1898	5.70E-196	74	100	GLUTAMYL-TRNA SYNTHETASE (GLUTAMATE--TRNA LIGASE) (GLURS) [Bacillus subtilis]
1291	Bt1Gc1638	Bt1G2815	7141-7984	g549024	1258	1260	2.30E-128	82	60	CYSTEINYL-TRNA SYNTHETASE (CYSTEINE--TRNA LIGASE) (CYSRS) [Bacillus subtilis]
1292	Bt1Gc1634	Bt1G2816	622-1	g1763079	452	433	9.90E-41	48	56	(U69493) Phn W [Salmonella typhimurium]
1292	Bt1Gc1634	Bt1G2817	1432-1048	g2623262	603	628	2.20E-61	98	99	(AF030979) phosphonoacetaldehyde hydrolase [Bacillus cereus]
1292	Bt1Gc1634	Bt1G2818	4517-1454	g1469285	356	448	6.10E-42	25	80	(U05042) afuB gene product [Actinobacillus pleuropneumoniae]
1292	Bt1Gc1634	Bt1G2819	3178-1475	g3319739	193	288	1.00E-22	22	100	(AL031035) ABC transporter integral membrane protein [Streptomyces coelicolor]
1292	Bt1Gc1634	Bt1G2820	4231-2116	g3341854	628	596	5.30E-58	48	71	(AF077856) PotA [Actinobacillus actinomycetemcomitans]
1292	Bt1Gc1634	Bt1G2821	5274-4249	g1469286	361	415	8.00E-39	30	100	(U05042) afuA gene product [Actinobacillus pleuropneumoniae]
1292	Bt1Gc1634	Bt1G2822	7566-6217	g2633721	1396	1046	1.10E-105	60	100	(Z99111) similar to Na+-transporting ATP synthase [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1292	Bt1Gc1634	Bt1G2823	9964-8879	g160475	1419	1374	1.90E-140	70	100	(AF109909) PHA synthase PhaC [Bacillus megaterium]
1293	Bt1Gc1635	Bt1G2824	105-1	g2635899	124	127	2.60E-08	60	17	(Z99121) yvbH [Bacillus subtilis]
1293	Bt1Gc1635	Bt1G2825	2381-3652	g2829689	1577	1374	1.90E-140	71	100	TRIGGER FACTOR (TF) (VEGETATIVE PROTEIN 2) (VEG2) [Bacillus subtilis]
1293	Bt1Gc1635	Bt1G2826	3923-5179	g1296452	1835	1775	6.10E-183	87	100	(X95306) ClpX protein [Bacillus subtilis]
1293	Bt1Gc1635	Bt1G2827	5289-6942	g1708857	2089	2040	5.10E-211	72	100	ATP-DEPENDENT PROTEASE LA HOMOLOG [Bacillus subtilis]
1293	Bt1Gc1635	Bt1G2828	7148-9469	g585415	2995	2668	1.40E-277	73	100	ATP-DEPENDENT PROTEASE LA 1 [Bacillus subtilis]
1293	Bt1Gc1635	Bt1G2829	9469-9757	g586754	378	397	6.50E-37	77	49	HYPOTHETICAL GTP-BINDING PROTEIN IN LONA-HEMA INTERGENIC REGION (ORFX) [Bacillus subtilis]
1294	Bt1Gc1639	Bt1G2830	4796-5794	g2443247	507	486	2.40E-46	33	100	(D86417) YfmD [Bacillus subtilis]
1294	Bt1Gc1639	Bt1G2831	5784-6581	g2443249	576	562	2.10E-54	45	100	(D86417) YfmF [Bacillus subtilis]
1294	Bt1Gc1639	Bt1G2832	8061-9242	g732334	1318	1125	4.60E-114	64	100	HYPOTHETICAL 44.4 KD PROTEIN IN EPR-GALK INTERGENIC REGION [Bacillus subtilis]
1294	Bt1Gc1639	Bt1G2833	9842-10516	g1175624	750	763	1.10E-75	65	100	HYPOTHETICAL 25.8 KD SENSORY TRANSDUCTION PROTEIN (ORF11) [Bacillus subtilis]
1294	Bt1Gc1639	Bt1G2834	10424-11470	g4481748	314	370	4.70E-34	27	100	(AF007865) BacS [Bacillus licheniformis]
1294	Bt1Gc1639	Bt1G2835	11646-12259	g3183506	628	669	9.70E-66	61	73	HYPOTHETICAL 31.7 KD PROTEIN IN GLTP-CWLJ INTERGENIC REGION (ORF13) [Bacillus subtilis]
1295	Bt1Gc1640	Bt1G2836	1079-1	g3257222	464	540	4.60E-52	35	97	(AP000003) 381aa long hypothetical protein [Pyrococcus horikoshii]
1295	Bt1Gc1640	Bt1G2837	2127-1219	g2633175	741	786	3.90E-78	48	100	(Z99108) similar to cell-division inhibitor [Bacillus subtilis]
1295	Bt1Gc1640	Bt1G2838	2209-3012	g2633176	508	484	3.90E-46	40	100	(Z99108) yfhG [Bacillus subtilis]
1295	Bt1Gc1640	Bt1G2839	3038-3349	g2633177	246	268	3.00E-23	55	99	(Z99108) yfhH [Bacillus subtilis]
1295	Bt1Gc1640	Bt1G2840	5156-4179	g2633185	1006	921	1.90E-92	58	100	(Z99108) yfhP [Bacillus subtilis]
1295	Bt1Gc1640	Bt1G2841	5300-6409	g2633186	1142	1109	2.30E-112	58	100	(Z99108) similar to A/G-specific adenine glycosylase [Bacillus subtilis]
1295	Bt1Gc1640	Bt1G2842	6663-6442	g2633187	281	281	1.30E-24	62	99	(Z99108) yfhS [Bacillus subtilis]
1295	Bt1Gc1640	Bt1G2843	7792-8107	g1673391	398	413	1.30E-38	74	99	(Z82044) hypothetical 12.2 kd protein [Bacillus subtilis]
1295	Bt1Gc1640	Bt1G2844	8352-10116	g1673392	1905	1721	3.20E-177	60	100	(Z82044) unidentified transporter-ATP binding [Bacillus subtilis]
1295	Bt1Gc1640	Bt1G2845	11135-10170	g1673393	811	719	4.90E-71	50	90	(Z82044) hypothetical 40.7 kd protein [Bacillus subtilis]
1296	Bt1Gc1641	Bt1G2846	82-4639	g2633966	586	626	3.50E-60	71	15	(Z99112) chromosome segregation SMC protein homolog [Bacillus subtilis]
1296	Bt1Gc1641	Bt1G2847	1225-3058	g1418413	1137	1112	1.10E-112	44	100	(X98606) potential coding region [Clostridium difficile]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1296	Bt1Gc1641	Bt1G2848	3174-4640	g4322642	445	562	2.10E-54	37	88	(AF093246) NAD(P)H-dependent 2-cyclohexen-1-one reductase Ncr [Pseudomonas syringae pv. glycinea]
1296	Bt1Gc1641	Bt1G2849	4131-5699	g2577963	1223	1216	1.10E-123	45	100	(Y15254) YerD protein [Bacillus subtilis]
1297	Bt1Gc1642	Bt1G2850	893-1	g2634032	1204	1148	1.70E-116	81	80	(Z99112) nusA [Bacillus subtilis]
1297	Bt1Gc1642	Bt1G2851	1381-914	g418459	558	435	6.10E-41	69	99	HYPOTHETICAL 17.6 KD PROTEIN IN NUSA 5'REGION (P15A) (ORF1) [Bacillus subtilis]
1297	Bt1Gc1642	Bt1G2852	6019-1721	g118793	5710	5507	0	75	100	DNA POLYMERASE III, ALPHA CHAIN [Bacillus subtilis]
1297	Bt1Gc1642	Bt1G2853	7844-6153	g2634029	2086	2089	3.30E-216	69	100	(Z99112) polyl-tRNA synthetase [Bacillus subtilis]
1297	Bt1Gc1642	Bt1G2854	9216-7957	g2634028	1305	1264	8.70E-129	60	100	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
1297	Bt1Gc1642	Bt1G2855	10370-9211	g3913544	1272	1276	4.60E-130	68	100	1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE (DXP REDUCTOISOMERASE) [Bacillus subtilis]
1297	Bt1Gc1642	Bt1G2856	11194-10394	g3121841	750	681	5.20E-67	55	100	PHOSPHATIDATE CYTIDYL TRANSFERASE (CDP-DIGLYCERIDE SYNTHETASE) (CDP-DIGLYCERIDE PYROPHOSPHORYLASE) (CDP-DIACYLGLYCEROL SYNTHASE) (CDS) (CTP:PHOSPHATIDATE CYTIDYL TRANSFERASE) (CDP-DAG SYNTHASE) [Bacillus subtilis]
1297	Bt1Gc1642	Bt1G2857	11985-11206	g3183468	905	796	3.40E-79	65	100	HYPOTHETICAL 29.8 KD PROTEIN IN FRR-CDSA INTERGENIC REGION [Bacillus subtilis]
1297	Bt1Gc1642	Bt1G2858	12622-12068	g3122762	714	633	6.40E-62	75	99	RIBOSOME RECYCLING FACTOR (RIBOSOME RELEASING FACTOR) (RRF) (VEGETATIVE PROTEIN 12B) (VEG12B) [Bacillus subtilis]
1297	Bt1Gc1642	Bt1G2859	12980-12628	g2634023	510	525	1.80E-50	87	49	(Z99112) uridylylate kinase [Bacillus subtilis]
1298	Bt1Gc1644	Bt1G2860	3552-1163	g586876	240	284	9.50E-25	32	52	HYPOTHETICAL 47.7 KD PROTEIN IN METS-KSGA INTERGENIC REGION [Bacillus subtilis]
1298	Bt1Gc1644	Bt1G2861	431-3640	g3256896	258	309	1.30E-26	27	39	(AP000002) 739aa long hypothetical chemotaxis protein [Pyrococcus horikoshii]
1298	Bt1Gc1644	Bt1G2862	8116-6789	g2661710	712	711	3.50E-70	35	100	(AL009204) FAD-dependent oxidoreductase [Streptomyces coelicolor]
1298	Bt1Gc1644	Bt1G2863	9274-8078	g3581864	622	673	3.70E-66	36	100	(AL031541) hypothetical protein SCI35.31 [Streptomyces coelicolor]
1298	Bt1Gc1644	Bt1G2864	11284-9358	g2126777	2954	2928	4.00E-305	93	100	phospholipase C (EC 3.1.4.3) - Bacillus cereus (strain IAM 1208) [Bacillus cereus]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1298	Bt1Gc1644	Bt1G2865	11626-11911	g420808	266	236	7.50E-20	63	97	hypothetical 12K protein - Thermus aquaticus [Thermus aquaticus]
1299	Bt1Gc1643	Bt1G2866	692-1	g728788	804	824	3.70E-82	64	40	ACETYL-COENZYME A SYNTHETASE (ACETATE--COA LIGASE) (ACYL-ACTIVATING ENZYME) (ACETYL-COA SYNTHASE) [Bacillus subtilis]
1299	Bt1Gc1643	Bt1G2867	869-1498	g728799	794	752	1.60E-74	69	100	ACETOIN UTILIZATION ACUA PROTEIN [Bacillus subtilis]
1299	Bt1Gc1643	Bt1G2868	1519-2160	g728800	534	545	1.30E-52	49	100	ACETOIN UTILIZATION ACUB PROTEIN [Bacillus subtilis]
1299	Bt1Gc1643	Bt1G2869	2157-3320	g728801	1400	1306	3.10E-133	65	100	ACETOIN UTILIZATION ACUC PROTEIN [Bacillus subtilis]
1299	Bt1Gc1643	Bt1G2870	5561-4850	g4481749	485	526	1.40E-50	42	100	(AF007865) BacR [Bacillus licheniformis]
1299	Bt1Gc1643	Bt1G2871	7734-5337	g4104606	390	432	1.30E-40	32	76	(AF036967) putative histidine kinase [Lactobacillus sakei]
1300	Bt1Gc1645	Bt1G2872	2990-528	g121881	3337	3249	0	78	100	DNA GYRASE SUBUNIT A [Bacillus subtilis]
1300	Bt1Gc1645	Bt1G2873	4995-3085	g121887	2707	2489	1.30E-258	79	100	DNA GYRASE SUBUNIT B [Bacillus subtilis]
1300	Bt1Gc1645	Bt1G2874	6170-5061	g132246	1309	1180	6.90E-120	66	100	REC F PROTEIN [Bacillus subtilis]
1300	Bt1Gc1645	Bt1G2875	7668-6526	g118797	1414	1438	3.20E-147	75	100	DNA POLYMERASE III, BETA CHAIN [Bacillus subtilis]
1301	Bt1Gc1646	Bt1G2876	25-579	g1001205	286	312	6.60E-28	35	99	(D64003) hypothetical protein [Synechocystis sp.]
1301	Bt1Gc1646	Bt1G2877	8164-8931	g729326	277	312	6.60E-28	29	100	GLUCOSE 1-DEHYDROGENASE III (GLCDH-III) [Bacillus megaterium]
1301	Bt1Gc1646	Bt1G2878	9002-9932	g732393	990	876	1.10E-87	61	100	HYPOTHETICAL 34.3 KD PROTEIN IN BGLH-WAPA INTERGENIC REGION (ORF1) [Bacillus subtilis]
1301	Bt1Gc1646	Bt1G2879	10177-10627	g1881269	320	302	7.60E-27	43	94	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
1302	Bt1Gc1649	Bt1G2880	1902-1	g3036999	348	381	2.50E-37	34	71	(AF019629) putative fimbria-associated protein [Actinomyces naeslundii]
1302	Bt1Gc1649	Bt1G2881	7096-7623	g1788724	180	205	1.40E-16	32	71	(AE000326) putative 2-component transcriptional regulator [Escherichia coli]
1302	Bt1Gc1649	Bt1G2882	8284-9207	g984656	680	701	4.00E-69	43	100	(D25292) ORF3 [Salmonella typhimurium]
1302	Bt1Gc1649	Bt1G2883	9339-9731	g1763704	380	151	6.80E-16	60	99	(Z83337) highly similar to E. coli large conductance mechanosensitive channel protein [Bacillus subtilis]
1302	Bt1Gc1649	Bt1G2884	12149-11225	g115950	1180	1191	4.70E-121	75	92	GLUCOSE-RESISTANCE AMYLASE REGULATOR (CATABOLITE CONTROL PROTEIN) [Bacillus subtilis]
1303	Bt1Gc1647	Bt1G2885	1-353	g115692	262	285	4.80E-25	43	57	CHLORAMPHENICOL ACETYLTRANSFERASE [Plasmid NR79]
1303	Bt1Gc1647	Bt1G2886	6820-1254	g2622173	356	421	1.40E-55	26	75	(AE000878) putative membrane protein [Methanobacterium

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
										thermoautotrophicum]
1303	Bt1Gc1647	Bt1G2887	7991-8743	g2635679	457	496	2.10E-47	42	100	(Z99120) similar to sepiapterin reductase [Bacillus subtilis]
1304	Bt1Gc1648	Bt1G2888	145-643	g2226153	362	387	7.40E-36	49	99	(Y14080) hypothetical protein [Bacillus subtilis]
1304	Bt1Gc1648	Bt1G2889	610-1648	g2226152	1394	1355	2.00E-138	76	100	(Y14080) hypothetical protein [Bacillus subtilis]
1304	Bt1Gc1648	Bt1G2890	1978-3531	g2808436	642	783	8.10E-78	35	100	(AL021309) fadD13 [Mycobacterium tuberculosis]
1304	Bt1Gc1648	Bt1G2891	5064-6536	g2326750	1465	988	1.50E-99	57	100	(Z98268) hypothetical protein Rv1707 [Mycobacterium tuberculosis]
1304	Bt1Gc1648	Bt1G2892	6865-8397	g267023	281	372	2.90E-34	23	100	STAGE V SPORULATION PROTEIN B [Bacillus subtilis]
1304	Bt1Gc1648	Bt1G2893	6847-8144	g586884	416	481	8.10E-46	29	81	HYPOTHETICAL 57.4 KD PROTEIN IN MFD-DIVIC INTERGENIC REGION [Bacillus subtilis]
1304	Bt1Gc1648	Bt1G2894	8532-9587	g1652052	245	353	3.00E-32	28	100	(D90902) hypothetical protein [Synechocystis sp.]
1304	Bt1Gc1648	Bt1G2895	8773-10408	g121838	115	197	2.00E-14	22	24	ENDOGLUCANASE PRECURSOR (ENDO-1,4-BETA-GLUCANASE) (ALKALINE CELLULASE) [Bacillus sp.]
1304	Bt1Gc1648	Bt1G2896	8337-11407	g4218533	274	410	8.30E-38	36	39	(AJ010312) endo-beta-N-acetylglucosaminidase [Streptococcus pneumoniae]
1305	Bt1Gc1651	Bt1G2897	1201-1	g2293313	1329	958	2.30E-96	65	93	(AF008220) YtiP [Bacillus subtilis]
1305	Bt1Gc1651	Bt1G2898	1404-2267	g3123307	974	1002	5.00E-101	68	100	HYPOTHETICAL OXIDOREDUCTASE IN APRE-COMK INTERGENIC REGION (ORFX) [Bacillus subtilis]
1305	Bt1Gc1651	Bt1G2899	2340-2897	g1731066	218	291	1.10E-25	34	99	HYPOTHETICAL 19.7 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1305	Bt1Gc1651	Bt1G2900	4647-4238	g2635860	385	411	2.10E-38	58	99	(Z99121) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1305	Bt1Gc1651	Bt1G2901	5069-4830	g2226143	222	225	1.10E-18	49	99	(Y14079) hypothetical protein [Bacillus subtilis]
1305	Bt1Gc1651	Bt1G2902	6860-5772	g3290176	775	691	4.50E-68	39	100	(AF067645) spore germination protein GerIB [Bacillus cereus]
1305	Bt1Gc1651	Bt1G2903	7959-6880	g3290177	356	361	4.20E-33	30	100	(AF067645) spore germination protein GerIC [Bacillus cereus]
1305	Bt1Gc1651	Bt1G2904	10964-7959	g3290175	860	730	3.30E-72	40	78	(AF067645) spore germination protein GerIA [Bacillus cereus]
1305	Bt1Gc1651	Bt1G2905	10452-9646	g2098612	342	433	9.90E-41	38	100	(U66614) putative transposase [Marinococcus halophilus]
1305	Bt1Gc1651	Bt1G2906	12176-11080	g1177016	1258	1111	1.40E-112	71	91	HYPOTHETICAL 43.7 KD PROTEIN IN KATB 3'REGION [Bacillus subtilis]
1306	Bt1Gc1653	Bt1G2907	10231-12365	g3150046	2253	2252	1.70E-233	63	82	(AF016634) ClpB chaperone homolog [Lactococcus lactis subsp. cremoris]
1307	Bt1Gc1657	Bt1G2908	999-1	g586848	1383	1339	9.80E-137	81	91	HYPOTHETICAL 40.1 KD GTP-BINDING PROTEIN IN

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1307	Bt1Gc1657	Bt1G2909	2474-3088	g586851	484	501	6.20E-48	48	100	RPSF-SPO0J INTERGENIC REGION [Bacillus subtilis]
1307	Bt1Gc1657	Bt1G2910	4696-3938	g586852	1061	1061	2.80E-107	81	100	HYPOTHETICAL 22.5 KD PROTEIN IN RPSF-SPO0J INTERGENIC REGION [Bacillus subtilis]
1307	Bt1Gc1657	Bt1G2911	5759-4896	g586854	940	969	1.60E-97	67	100	SOJ PROTEIN [Bacillus subtilis]
1307	Bt1Gc1657	Bt1G2912	6584-5868	g121190	887	796	3.40E-79	70	100	HYPOTHETICAL 32.8 KD PROTEIN IN SPO0J-GIDB INTERGENIC REGION [Bacillus subtilis]
1307	Bt1Gc1657	Bt1G2913	8495-6612	g121187	2703	2703	2.80E-281	82	100	GLUCOSE INHIBITED DIVISION PROTEIN B [Bacillus subtilis]
1307	Bt1Gc1657	Bt1G2914	9912-8545	g135725	1709	1604	8.10E-165	73	100	GLUCOSE INHIBITED DIVISION PROTEIN A [Bacillus subtilis]
1308	Bt1Gc1652	Bt1G2915	1-716	g1834395	1099	999	1.00E-100	96	86	POSSIBLE THIOPHENE AND FURAN OXIDATION PROTEIN THDF [Bacillus subtilis]
1308	Bt1Gc1652	Bt1G2917	10422-9088	g3322685	728	792	9.00E-79	40	100	(X67138) flagellin [Bacillus thuringiensis]
1308	Bt1Gc1652	Bt1G2918	12123-11098	g1706844	459	493	4.40E-47	35	100	(AE001218) flagellum-specific ATP synthase (fliI) [Treponema pallidum]
1309	Bt1Gc1654	Bt1G2919	1065-139	g127029	812	832	5.20E-83	50	100	FLAGELLAR MOTOR SWITCH PROTEIN FLIG [Borrelia burgdorferi]
1309	Bt1Gc1654	Bt1G2920	2374-1073	g1835113	1006	1076	7.20E-109	48	100	HOMOSERINE O-SUCCINYLTRANSFERASE (HOMOSERINE O-TRANSUCCINYLAASE) [Escherichia coli]
1309	Bt1Gc1654	Bt1G2921	4836-4145	g2116760	794	718	6.30E-71	64	97	O-acetylhomoserine sulphydrylase [Leptospira meyeri]
1309	Bt1Gc1654	Bt1G2922	4983-5642	g465600	410	431	1.60E-40	42	100	(D86418) YfnB [Bacillus subtilis]
1309	Bt1Gc1654	Bt1G2923	6303-5680	g2619052	766	686	1.50E-67	67	100	HYPOTHETICAL 25.3 KD PROTEIN IN NFO-FRUA INTERGENIC REGION [Escherichia coli]
1309	Bt1Gc1654	Bt1G2924	7168-6483	g4753864	460	501	6.20E-48	44	100	(AF027868) acyl carrier protein phosphodiesterase [Bacillus subtilis]
1309	Bt1Gc1654	Bt1G2925	8731-7142	g2226140	778	853	3.10E-85	48	100	putative two-component system regulator [Streptomyces coelicolor]
1309	Bt1Gc1654	Bt1G2926	8869-9678	g732357	896	910	2.80E-91	66	100	(Y14079) hypothetical protein [Bacillus subtilis]
1309	Bt1Gc1654	Bt1G2927	10838-9774	g1724005	391	343	3.40E-31	31	100	PHOSPHOMETHYL PYRIMIDINE KINASE (HMP-PHOSPHATE KINASE) (HMP-P KINASE) [Bacillus subtilis]
1309	Bt1Gc1654	Bt1G2927	10838-9774	g1724005	391	343	3.40E-31	31	100	HYPOTHETICAL 40.7 KD PROTEIN IN CSPB-GLPP INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1309	Bt1Gc1654	Bt1G2928	11187-10987	g732103	131	138	1.80E-09	37	99	HYPOTHETICAL 7.7 KD PROTEIN IN MRR-TSR INTERGENIC REGION (F67) [Escherichia coli]
1309	Bt1Gc1654	Bt1G2929	13243-11163	g2506997	2207	1795	4.70E-185	60	99	CARBON STARVATION PROTEIN A [Escherichia coli]
1310	Bt1Gc1650	Bt1G2930	269-2182	g2633756	1789	1679	9.10E-173	55	100	(Z99111) similar to heavy metal-transporting ATPase [Bacillus subtilis]
1310	Bt1Gc1650	Bt1G2931	3248-2382	g4584097	1398	925	7.30E-93	95	100	(AJ010131) yfkH [Bacillus cereus]
1310	Bt1Gc1650	Bt1G2932	3447-3819	g4584098	577	200	4.90E-16	95	99	(AJ010131) hypothetical protein [Bacillus cereus]
1310	Bt1Gc1650	Bt1G2933	4288-3827	g2626822	480	490	9.10E-47	59	99	(D83967) Yfkj [Bacillus subtilis]
1310	Bt1Gc1650	Bt1G2934	4882-7560	g2337795	1743	1112	3.00E-185	46	100	(Y13937) putative PacL protein [Bacillus subtilis]
1310	Bt1Gc1650	Bt1G2935	8845-7751	g2415745	1087	1132	8.40E-115	59	100	(AB000617) YceH [Bacillus subtilis]
1311	Bt1Gc1658	Bt1G2936	2744-1924	g3913256	317	315	3.20E-28	36	62	SENSOR PROTEIN CIAH [Streptococcus pneumoniae]
1311	Bt1Gc1658	Bt1G2937	9450-8338	g2492593	740	660	8.80E-65	47	100	PROBABLE ABC TRANSPORTER ATP-BINDING PROTEIN Y4OS [Rhizobium sp. NGR234]
1311	Bt1Gc1658	Bt1G2938	11325-10723	g2492902	342	348	1.00E-31	38	100	HYPOTHETICAL 21.7 KD PROTEIN IN BETT-PRPR INTERGENIC REGION [Escherichia coli]
1311	Bt1Gc1658	Bt1G2939	13146-13639	g1177013	604	553	1.90E-53	70	38	HYPOTHETICAL 45.5 KD PROTEIN IN BGLS-KATB INTERGENIC REGION [Bacillus subtilis]
1312	Bt1Gc1660	Bt1G2940	178-1	g224807	107	132	4.30E-08	41	17	gene tnpA [Staphylococcus aureus]
1312	Bt1Gc1660	Bt1G2941	841-1815	g2633127	1354	1303	6.40E-133	79	100	(Z99108) similar to hypothetical proteins [Bacillus subtilis]
1312	Bt1Gc1660	Bt1G2942	2187-2984	g465586	245	337	1.50E-30	42	99	HYPOTHETICAL 18.0 KD PROTEIN IN MOLR-BGLX INTERGENIC REGION [Escherichia coli]
1312	Bt1Gc1660	Bt1G2943	2795-3565	g3257120	122	261	1.70E-22	25	100	(AP000003) 230aa long hypothetical protein [Pyrococcus horikoshii]
1312	Bt1Gc1660	Bt1G2944	5522-4434	g2632019	545	551	3.10E-53	35	100	(AJ002571) YkfB [Bacillus subtilis]
1312	Bt1Gc1660	Bt1G2945	8299-6903	g1731094	841	678	1.10E-66	38	100	HYPOTHETICAL NA+/H+ ANTIPORTER IN ANSB-SPOIIM INTERGENIC REGION [Bacillus subtilis]
1312	Bt1Gc1660	Bt1G2946	8584-9726	g1323764	987	997	1.70E-100	50	100	(U47860) amidohydrolase AmhX from B. subtilis [Bacillus subtilis]
1312	Bt1Gc1660	Bt1G2947	12077-10552	g118589	2239	1978	1.90E-204	86	98	NADH DEHYDROGENASE (ALKYL HYDROPEROXIDE REDUCTASE) [Bacillus alcalophilus]
1313	Bt1Gc1659	Bt1G2948	1-691	g1177013	765	725	1.10E-71	65	54	HYPOTHETICAL 45.5 KD PROTEIN IN BGLS-KATB INTERGENIC REGION [Bacillus subtilis]
1313	Bt1Gc1659	Bt1G2949	1422-742	g2443239	444	444	6.80E-42	41	100	(D86417) YfiQ [Bacillus subtilis]
1313	Bt1Gc1659	Bt1G2950	3813-1427	g1705888	756	823	4.70E-82	34	100	SENSOR KINASE CITa [Klebsiella pneumoniae]
1313	Bt1Gc1659	Bt1G2951	5171-3189	g730002	1440	1227	7.20E-125	44	100	METHYL-ACCEPTING CHEMOTAXIS PROTEIN MCPA

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1313	Bt1Gc1659	Bt1G2952	6981-5413	g3915560	1584	1109	2.30E-112	64	97	(H1) [Bacillus subtilis] HYPOTHETICAL 56.0 KD PROTEIN IN GLGB-GBSB INTERGENIC REGION [Bacillus subtilis]
1313	Bt1Gc1659	Bt1G2953	7542-7016	g2635586	271	188	4.90E-17	37	99	(Z99119) yuaF [Bacillus subtilis]
1313	Bt1Gc1659	Bt1G2954	10755-6797	g4760825	2056	2183	3.60E-226	43	92	(D87215) collagenase [Clostridium histolyticum]
1314	Bt1Gc1655	Bt1G2955	1-752	g4768683	190	330	8.20E-30	34	72	(AF115385) putative ferric enterobactin periplasmic binding protein [Neisseria gonorrhoeae]
1314	Bt1Gc1655	Bt1G2956	821-1770	g1107528	632	462	8.40E-44	39	100	(X88849) tlg start [Campylobacter coli]
1314	Bt1Gc1655	Bt1G2957	1899-2843	g1805451	497	467	2.50E-44	33	100	(D50453) homologue of ferric anguibactin transport system permerase protein FatC of V. anguillarum [Bacillus subtilis]
1314	Bt1Gc1655	Bt1G2958	2840-3595	g1805452	596	602	1.20E-58	46	100	(D50453) homologue of iron dicitrate transport ATP-binding protein FecE of E. coli [Bacillus subtilis]
1314	Bt1Gc1655	Bt1G2959	4059-5511	g3334367	2327	2373	2.60E-246	94	100	HEMOLYSIN PRECURSOR [Bacillus cereus]
1314	Bt1Gc1655	Bt1G2960	8219-8626	g2460264	370	250	2.40E-21	51	99	(AF020807) Mig-7 [Salmonella typhimurium]
1314	Bt1Gc1655	Bt1G2961	8769-9323	g1742583	642	616	4.00E-60	63	99	(D90799) Spermidine N1-acetyltransferase (EC 2.3.1.57) (Diamine acetyltransferase) (SAT). [Escherichia coli]
1314	Bt1Gc1655	Bt1G2962	9725-11269	g2443256	1083	1039	6.00E-105	42	100	(D86417) YfmM [Bacillus subtilis]
1314	Bt1Gc1655	Bt1G2963	11275-12479	g1800301	532	451	1.20E-42	32	100	(U83667) macrolide-efflux determinant [Streptococcus pneumoniae]
1315	Bt1Gc1656	Bt1G2964	1-604	g2626830	563	595	6.80E-58	56	36	(D83967) TreC [Bacillus subtilis]
1315	Bt1Gc1656	Bt1G2965	1282-917	g2072375	589	589	2.90E-57	96	99	(Y09254) hypothetical protein [Bacillus cereus]
1315	Bt1Gc1656	Bt1G2966	4383-2791	g1805440	865	886	9.90E-89	35	100	(D50453) spore germination protein GerKA [Bacillus subtilis]
1315	Bt1Gc1656	Bt1G2967	4509-5942	g2577967	335	459	1.70E-43	30	100	(Y15254) YerI protein [Bacillus subtilis]
1315	Bt1Gc1656	Bt1G2968	5649-7106	g2635246	1757	1727	7.50E-178	68	100	(Z99118) similar to sodium/proton-dependent alanine carrier protein [Bacillus subtilis]
1315	Bt1Gc1656	Bt1G2969	7245-7970	g2635188	972	868	8.00E-87	74	100	(Z99117) glutamine ABC transporter (ATP-binding protein) [Bacillus subtilis]
1316	Bt1Gc1662	Bt1G2970	687-899	g2707959	183	193	2.70E-15	49	99	(Z97974) Tec protein [Lactobacillus bacteriophage phi adh]
1316	Bt1Gc1662	Bt1G2971	2946-3806	g1926335	165	291	1.10E-25	28	100	(X98106) zinc finger protein [Bacteriophage phig1e]
1316	Bt1Gc1662	Bt1G2972	9654-9938	g2634222	146	164	3.20E-12	38	99	(Z99113) yoeC [Bacillus subtilis]
1316	Bt1Gc1662	Bt1G2973	9651-10196	g2687738	205	246	6.50E-21	29	99	(AJ222725) hypothetical protein [Lactobacillus helveticus]
1317	Bt1Gc1661	Bt1G2974	1-368	g2497392	575	592	1.40E-57	93	49	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
1317	Bt1Gc1661	Bt1G2975	3965-2943	g1787450	624	562	2.10E-69	40	100	(AE000218) putative dihydroxyacetone kinase (EC 2.7.1.2)



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1317	Bt1Gc1661	Bt1G2976	4670-5701	g1787450	463	421	1.90E-52	35	100	[Escherichia coli] putative dihydroxyacetone kinase (EC 2.7.1.2) [Escherichia coli]
1317	Bt1Gc1661	Bt1G2977	6523-5918	g3183496	313	300	1.20E-26	33	100	HYPOTHETICAL 23.4 KD PROTEIN IN AAPA-SIGV INTERGENIC REGION [Bacillus subtilis]
1317	Bt1Gc1661	Bt1G2978	10089-11255	g1742730	419	380	4.10E-35	29	100	(D90809) Protein Aral precursor. [Escherichia coli]
1317	Bt1Gc1661	Bt1G2979	12204-11780	g1789981	387	406	7.20E-38	52	50	(AE000433) IS150 putative transposase [Escherichia coli]
1318	Bt1Gc1663	Bt1G2980	392-1264	g3122275	803	825	2.90E-82	52	100	PUTATIVE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (TRANSAMINASE B) (BCAT) [Archaeoglobus fulgidus]
1318	Bt1Gc1663	Bt1G2981	1580-3292	g1770066	1908	1880	4.60E-194	64	100	(Z75208) acetolactate synthase large subunit [Bacillus subtilis]
1318	Bt1Gc1663	Bt1G2982	3295-3804	g1170549	524	492	5.60E-47	63	99	ACETOLACTATE SYNTHASE SMALL SUBUNIT (AHAS) (ACETOHYDROXY-ACID SYNTHASE SMALL SUBUNIT) (ALS) [Bacillus subtilis]
1318	Bt1Gc1663	Bt1G2983	3831-4856	g585314	1235	1250	2.60E-127	69	100	KETOL-ACID REDUCTOISOMERASE (ACETOHYDROXY-ACID ISOMEROREDUCTASE) (ALPHA-KETO-BETA-HYDROXYLACIL REDUCTOISOMERASE) [Bacillus subtilis]
1318	Bt1Gc1663	Bt1G2984	4843-6379	g2492784	1410	1441	1.50E-147	56	100	2-ISOPROPYLMALATE SYNTHASE (ALPHA-ISOPROPYLMALATE SYNTHASE) (ALPHA-IPM SYNTHETASE) [Bacillus subtilis]
1318	Bt1Gc1663	Bt1G2985	6441-7513	g1770070	927	987	2.00E-99	55	100	(Z75208) 3-isopropylmalate dehydrogenase [Bacillus subtilis]
1318	Bt1Gc1663	Bt1G2986	7487-9799	g3122355	1487	1208	2.10E-173	48	100	3-ISOPROPYLMALATE DEHYDRATASE (ISOPROPYLMALATE ISOMERASE) (ALPHA-IPM ISOMERASE) (IPMI) [Schizosaccharomyces pombe]
1318	Bt1Gc1663	Bt1G2987	9774-11949	g2618864	457	544	1.70E-52	33	89	(AF017113) histidyl-tRNA synthetase [Bacillus subtilis]
1318	Bt1Gc1663	Bt1G2988	11003-11644	g3122204	455	488	1.50E-46	44	100	ATP PHOSPHORIBOSYLTRANSFERASE [Bacillus subtilis]
1318	Bt1Gc1663	Bt1G2989	11635-12268	g3122231	402	457	2.80E-43	46	49	HISTIDINOL DEHYDROGENASE (HDH) [Methanococcus jannaschii]
1319	Bt1Gc1666	Bt1G2990	2305-1	g3123226	371	474	2.10E-44	28	63	INTERNALIN A PRECURSOR []
1319	Bt1Gc1666	Bt1G2991	3660-2478	g3915989	1669	1502	5.20E-154	79	100	HYPOTHETICAL 45.3 KD PROTEIN IN PRKA-CSPB INTERGENIC REGION (ORF4) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1319	Bt1Gc1666	Bt1G2992	6008-4116	g730399	2938	2938	3.50E-306	88	100	PRKA PROTEIN [Bacillus subtilis]
1319	Bt1Gc1666	Bt1G2993	9541-6407	g3659614	589	645	3.40E-63	32	58	(AF052517) c-di-GMP phosphodiesterase A [Acetobacter xylinus]
1319	Bt1Gc1666	Bt1G2994	9080-8610	g1381681	619	620	1.50E-60	70	99	(U58864) CspR [Bacillus subtilis]
1319	Bt1Gc1666	Bt1G2995	10023-9130	g418441	606	609	2.20E-59	42	100	HYPOTHETICAL 35.5 KD PROTEIN IN GLDA 3-REGION (ORF4) [Bacillus stearothermophilus]
1319	Bt1Gc1666	Bt1G2996	11474-10060	g1903044	1401	1290	1.50E-131	64	99	(Z93102) hypothetical 48.5 kd protein [Bacillus subtilis]
1320	Bt1Gc1664	Bt1G2997	1-411	g1351072	317	358	1.00E-31	49	16	S-LAYER PROTEIN PRECURSOR (SURFACE LAYER PROTEIN) [Bacillus licheniformis]
1320	Bt1Gc1664	Bt1G2998	712-1370	g2634853	235	325	2.80E-29	39	100	(Z99116) alternate gene name: yqil; similar to N-acetylmuramoyl-L-alanine amidase [Bacillus subtilis]
1320	Bt1Gc1664	Bt1G2999	7540-8484	g4154809	574	331	4.40E-57	39	100	(AE001465) DIPEPTIDE TRANSPORT SYSTEM PERMEASE PROTEIN [Helicobacter pylori J99]
1320	Bt1Gc1664	Bt1G3000	8440-9336	g1168474	623	481	8.10E-46	44	100	OLIGOPEPTIDE TRANSPORT PERMEASE PROTEIN APPC [Bacillus subtilis]
1320	Bt1Gc1664	Bt1G3001	9378-11166	g1723073	1230	1225	1.20E-124	43	100	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CY50.01 [Mycobacterium tuberculosis]
1321	Bt1Gc1665	Bt1G3003	3970-6284	g2149596	2736	2597	4.80E-270	91	100	(U89796) chitinase [Bacillus thuringiensis]
1321	Bt1Gc1665	Bt1G3004	7393-6641	g1477569	440	481	8.10E-46	39	100	(U50335) ORF3 [Mycobacterium smegmatis]
1321	Bt1Gc1665	Bt1G3005	8874-7864	g1175228	173	262	1.30E-22	32	52	PUTATIVE THIAMINE BIOSYNTHESIS PROTEIN HI0357 [Haemophilus influenzae Rd]
1321	Bt1Gc1665	Bt1G3006	9137-8403	g2501390	405	405	9.20E-38	33	100	HYPOTHETICAL ABC TRANSPORTER PERMEASE PROTEIN HI0355 [Haemophilus influenzae Rd]
1321	Bt1Gc1665	Bt1G3007	9441-9154	g3025318	232	223	1.80E-18	46	99	HYPOTHETICAL 10.7 KD PROTEIN IN VIRR 5'REGION (ORF2) [Clostridium perfringens]
1322	Bt1Gc1669	Bt1G3008	3267-2725	g732321	554	548	6.50E-53	66	99	HYPOTHETICAL 22.0 KD PROTEIN IN FLIT-SECA INTERGENIC REGION [Bacillus subtilis]
1322	Bt1Gc1669	Bt1G3009	3786-3592	g2493763	341	341	5.60E-31	100	98	COLD SHOCK-LIKE PROTEIN CSPC [Bacillus cereus]
1322	Bt1Gc1669	Bt1G3010	4617-3916	g729159	366	467	2.50E-44	45	100	COMF OPERON PROTEIN 3 [Bacillus subtilis]
1322	Bt1Gc1669	Bt1G3011	6903-4620	g729157	1060	1130	1.40E-114	49	100	COMF OPERON PROTEIN 1 [Bacillus subtilis]
1322	Bt1Gc1669	Bt1G3012	6446-6094	g2226145	240	304	1.00E-26	47	24	(Y14079) hypothetical protein [Bacillus subtilis]
1322	Bt1Gc1669	Bt1G3013	7371-4824	g129375	219	122	1.40E-05	24	60	P54 PROTEIN PRECURSOR []
1322	Bt1Gc1669	Bt1G3014	7963-7511	g2635881	259	326	2.20E-29	45	64	(Z99121) similar to hypothetical proteins [Bacillus subtilis]
1322	Bt1Gc1669	Bt1G3015	8984-8145	g729311	906	874	1.80E-87	63	100	DEGV PROTEIN [Bacillus subtilis]
1322	Bt1Gc1669	Bt1G3016	9221-9867	g3123311	846	857	1.20E-85	73	100	HYPOTHETICAL 24.8 KD PROTEIN IN DEGS-TAGO

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1322	Bt1Gc1669	Bt1G3017	9939-11080	g1762327	666	648	1.60E-63	40	100	INTERGENIC REGION [Bacillus subtilis] (U56901) putative transcriptional regulator [Bacillus subtilis]
1322	Bt1Gc1669	Bt1G3018	12213-11069	g732324	1368	1390	3.90E-142	70	100	PUTATIVE UDP-N-ACETYLGLUCOSAMINE 2-EPIMERASE (UDP-GLCNAC-2-EPIMERASE) [Bacillus subtilis]
1322	Bt1Gc1669	Bt1G3019	13616-12300	g1894744	1210	1229	4.40E-125	54	100	product similar to Pseudomonas aeruginosa GDP-mannose 6-dehydrogenase protein [Bacillus subtilis]
1322	Bt1Gc1669	Bt1G3020	14146-15216	g2582651	1140	765	6.50E-76	63	100	(AJ004803) putative undecaprenyl-phosphate N-acetylglucosaminyltransferase [Bacillus subtilis]
1322	Bt1Gc1669	Bt1G3021	15549-16135	g2226222	280	372	2.90E-34	36	72	(Y14082) hypothetical protein [Bacillus subtilis]
1323	Bt1Gc1670	Bt1G3022	2597-996	g2293232	1795	1806	3.20E-186	64	100	(AF008220) YtcI [Bacillus subtilis]
1323	Bt1Gc1670	Bt1G3023	2977-2783	g134231	318	318	1.50E-28	100	98	SMALL, ACID-SOLUBLE SPORE PROTEIN 2 (SASP) [Bacillus cereus]
1323	Bt1Gc1670	Bt1G3024	4269-3196	g3915122	1049	895	1.10E-89	56	100	PROBABLE THIAMIN BIOSYNTHESIS PROTEIN THII [Bacillus subtilis]
1323	Bt1Gc1670	Bt1G3025	5422-4283	g2293229	920	895	1.10E-89	48	100	(AF008220) Nifs2 [Bacillus subtilis]
1323	Bt1Gc1670	Bt1G3026	7501-5819	g2293228	957	964	5.30E-97	36	100	(AF008220) YtwP [Bacillus subtilis]
1323	Bt1Gc1670	Bt1G3027	8563-6993	g732339	179	279	2.10E-24	25	100	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN THIK-EPR INTERGENIC REGION [Bacillus subtilis]
1323	Bt1Gc1670	Bt1G3028	8688-9650	g1934659	689	585	7.80E-57	39	100	(U93876) hypothetical protein YrdR [Bacillus subtilis]
1323	Bt1Gc1670	Bt1G3029	10285-9667	g2293227	481	499	1.00E-47	45	100	(AF008220) putative transcription regulator [Bacillus subtilis]
1323	Bt1Gc1670	Bt1G3030	10400-10948	g3915450	349	383	2.00E-35	42	99	HYPOTHETICAL 20.3 KD PROTEIN IN PRC-PPHA INTERGENIC REGION [Escherichia coli]
1324	Bt1Gc1671	Bt1G3031	646-2561	g4126632	338	445	5.30E-42	54	53	(AB016282) ORF46 [bacteriophage phi-105]
1324	Bt1Gc1671	Bt1G3032	1144-1674	g538957	125	221	8.10E-17	34	15	endo-1,4-beta-xylanase (EC 3.2.1.8) precursor - Thermoanaerobacterium saccharolyticum (strain B6A-RU) [Thermoanaerobacterium saccharolyticum]
1324	Bt1Gc1671	Bt1G3033	5976-5097	g3688821	699	717	8.00E-71	50	100	(AF084104) hypothetical protein [Bacillus firmus]
1324	Bt1Gc1671	Bt1G3034	6074-6693	g4376792	318*	394	1.30E-36	45	99	(AE001635) 3-methyladenine DNA glycosylase [Chlamydia pneumoniae]
1324	Bt1Gc1671	Bt1G3035	7875-7111	g543795	700	720	3.80E-71	50	100	ALPHA-ACETOLACTATE DECARBOXYLASE [Bacillus subtilis]
1324	Bt1Gc1671	Bt1G3036	9577-7869	g1929340	1888	1833	4.40E-189	63	100	(Z93767) alpha-acetolactate synthase protein, AlsS

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1324	Bt1Gc1671	Bt1G3037	11468-11326	g2226157	124	143	5.30E-10	56	40	[Bacillus subtilis] (Y14080) hypothetical protein [Bacillus subtilis]
1325	Bt1Gc1672	Bt1G3038	378-1	g2635185	457	438	2.90E-41	69	91	[Z99117] similar to hypothetical proteins [Bacillus subtilis]
1325	Bt1Gc1672	Bt1G3039	3367-731	g3122886	3462	3311	0	75	100	ALANYL-TRNA SYNTHETASE (ALANINE--TRNA LIGASE) (ALARS) [Bacillus subtilis]
1325	Bt1Gc1672	Bt1G3040	5324-4266	g2635187	848	598	3.30E-58	44	100	[Z99117] similar to hypothetical proteins [Bacillus subtilis]
1325	Bt1Gc1672	Bt1G3041	8558-6213	g2635193	2289	2240	3.30E-232	56	100	[Z99117] similar to conjugation transfer protein [Bacillus subtilis]
1325	Bt1Gc1672	Bt1G3042	9283-8666	g2635194	560	561	2.70E-54	53	100	[Z99117] similar to hypothetical proteins [Bacillus subtilis]
1325	Bt1Gc1672	Bt1G3043	10538-9432	g3122977	1509	1522	4.00E-156	75	100	PROBABLE TRNA (5-METHYLAMINOMETHYL-2-THIOURIDYLATE)-METHYLTRANSFERASE [Bacillus subtilis]
1325	Bt1Gc1672	Bt1G3044	11391-10666	g2635197	841	843	3.50E-84	64	70	[Z99117] similar to NifS protein homolog [Bacillus subtilis]
1326	Bt1Gc1673	Bt1G3045	3431-245	g1652490	661	771	1.50E-76	36	65	[D90905] hypothetical protein [Synechocystis sp.]
1326	Bt1Gc1673	Bt1G3046	2967-4490	g2952028	1063	1084	1.00E-109	41	100	[U88888] cardiolipin synthase [Bacillus firmus]
1326	Bt1Gc1673	Bt1G3047	4588-5505	g2634332	869	890	3.70E-89	58	100	[Z99114] similar to nitric-oxide reductase [Bacillus subtilis]
1326	Bt1Gc1673	Bt1G3048	5450-7398	g2634331	1155	1217	8.30E-124	40	100	[Z99114] yojO [Bacillus subtilis]
1326	Bt1Gc1673	Bt1G3049	9997-8016	g1458327	592	689	7.40E-68	43	59	[U64847] F08F3.4 gene product [Caenorhabditis elegans]
1326	Bt1Gc1673	Bt1G3050	10121-8931	g3256681	1057	1119	2.00E-113	55	100	[AP000001] 398aa long hypothetical 5-aminolevulinic acid synthase (8 amino-7-oxonanoate synthase) [Pyrococcus horikoshii]
1326	Bt1Gc1673	Bt1G3051	10338-11069	g1175329	254	300	1.20E-26	31	100	HYPOTHETICAL PROTEIN HI0912 [Haemophilus influenzae Rd]
1326	Bt1Gc1673	Bt1G3052	11918-11094	g2635790	1030	957	3.00E-96	71	100	[Z99120] similar to iron(III) dicitrate transport permease [Bacillus subtilis]
1327	Bt1Gc1674	Bt1G3053	1-1502	g2293198	1226	1147	2.20E-116	51	93	[AF008220] YtgP [Bacillus subtilis]
1327	Bt1Gc1674	Bt1G3054	1824-2291	g3915559	420	429	2.60E-40	54	99	HYPOTHETICAL 17.7 KD PROTEIN IN AMYX-OPUD INTERGENIC REGION [Bacillus subtilis]
1327	Bt1Gc1674	Bt1G3055	2607-2389	g2635486	291	291	1.10E-25	74	99	[Z99119] similar to transcriptional regulator (DeoR family) [Bacillus subtilis]
1327	Bt1Gc1674	Bt1G3056	2974-4368	g2293201	1400	1397	7.00E-143	58	100	[AF008220] putative peptidase [Bacillus subtilis]
1327	Bt1Gc1674	Bt1G3057	4472-6397	g1075814	317	398	5.10E-37	37	67	hypothetical protein 3 - Bacillus stearothermophilus [Bacillus stearothermophilus]
1327	Bt1Gc1674	Bt1G3058	5394-7373	g228654	1823	1774	7.80E-183	54	100	thermostable pullulanase [Bacillus stearothermophilus]
1327	Bt1Gc1674	Bt1G3059	7595-8403	g2293206	772	754	9.60E-75	54	100	[AF008220] YtmP [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aaf_nap Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1327	Bt1Gc1674	Bt1G3060	8680-8405	g2635475	279	1.40E-20	54	99	(Z99119) ytzH [Bacillus subtilis]
1327	Bt1Gc1674	Bt1G3061	8909-9547	g2293207	739	1.10E-66	65	100	(AF008220) YtmQ [Bacillus subtilis]
1327	Bt1Gc1674	Bt1G3062	9707-10474	g2293208	882	2.20E-82	62	100	(AF008220) YtmP [Bacillus subtilis]
1327	Bt1Gc1674	Bt1G3063	11036-10719	g2635471	171	7.40E-13	36	99	(Z99119) ytzB [Bacillus subtilis]
1327	Bt1Gc1674	Bt1G3064	11185-12255	g2293210	1446	3.50E-148	74	100	(AF008220) YtoP [Bacillus subtilis]
1327	Bt1Gc1674	Bt1G3065	12396-13051	g3915930	103	3.70E-18	27	100	HYPOTHETICAL PROTEIN MJ0261 [Methanococcus jannaschii]
1327	Bt1Gc1674	Bt1G3066	13132-13409	g2293211	314	3.20E-28	54	86	(AF008220) putative thioredoxin [Bacillus subtilis]
1328	Bt1Gc1676	Bt1G3067	7614-8411	g3386359	711	3.00E-71	53	100	(AF074855) RNA polymerase sigma B [Listeria monocytogenes]
1328	Bt1Gc1676	Bt1G3068	9054-11223	g1653076	164	2.10E-13	37	51	(D90910) regulatory components of sensory transduction system [Synecocystis sp.]
1328	Bt1Gc1676	Bt1G3069	8005-10219	g1945716	363	4.80E-41	31	68	(Z94043) hypothetical protein [Bacillus subtilis]
1328	Bt1Gc1676	Bt1G3070	11071-10292	g584925	227	7.60E-27	26	100	CHEMOTAXIS PROTEIN METHYLTRANSFERASE [Bacillus subtilis]
1328	Bt1Gc1676	Bt1G3071	16729-11142	g3294236	905	1.20E-99	35	40	(AL031031) putative sensory histidine kinase [Streptomyces coelicolor]
1328	Bt1Gc1676	Bt1G3072	15716-16030	g1787091	282	7.60E-27	55	99	(AE000188) orf, hypothetical protein [Escherichia coli]
1329	Bt1Gc1675	Bt1G3073	334-1	g1945709	244	1.00E-24	53	50	(Z94043) hypothetical protein [Bacillus subtilis]
1329	Bt1Gc1675	Bt1G3074	5171-6580	g1770026	1847	1.80E-183	75	100	(Z75208) hypothetical protein [Bacillus subtilis]
1329	Bt1Gc1675	Bt1G3075	6577-7908	g1770025	1622	8.20E-156	68	100	(Z75208) hypothetical protein [Bacillus subtilis]
1329	Bt1Gc1675	Bt1G3076	8374-9198	g2633768	521	6.50E-53	51	100	(Z99111) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1329	Bt1Gc1675	Bt1G3077	10868-9963	g1724014	726	3.30E-72	46	100	HYPOTHETICAL 33.7 KD PROTEIN IN CSPB-GLPP INTERGENIC REGION [Bacillus subtilis]
1329	Bt1Gc1675	Bt1G3078	13196-9890	g2104606	466	7.30E-55	28	62	(Z95398) unknown [Mycobacterium leprae]
1329	Bt1Gc1675	Bt1G3079	14449-13799	g322157	254	1.70E-22	31	100	hypothetical protein 2 - Bacillus megaterium [Bacillus megaterium]
1329	Bt1Gc1675	Bt1G3080	15552-14512	g1172561	564	6.90E-49	35	100	SPERMIDINE/PUTRESCINE-BINDING PERIPLASMIC PROTEIN 2 PRECURSOR (SPBP) []
1329	Bt1Gc1675	Bt1G3081	16391-15627	g1172564	526	1.90E-37	40	100	SPERMIDINE/PUTRESCINE TRANSPORT SYSTEM PERMEASE PROTEIN POTC []
1330	Bt1Gc1677	Bt1G3082	1-307	g3005554	93	5.60E-09	25	32	(AF047044) putative transposase [Anabaena PCC7120]
1330	Bt1Gc1677	Bt1G3083	10864-1	g1850913	130	1.10E-29	20	28	(L03534) myosin heavy chain [Entamoeba histolytica]
1330	Bt1Gc1677	Bt1G3084	4731-3283	g2444121	53	1.60E-10	37	7	(U88974) ORF42 [Streptococcus thermophilus temperate

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrq	NCBI gi description
1330	Bt1Gc1677	Bt1G3085	10492-7187	g2764873	347	541	3.60E-52	26	100	bacteriophage O1205]
1331	Bt1Gc1667	Bt1G3086	14-1381	g2126905	1731	1731	2.80E-178	72	100	(X97918) gene 18.1 [Bacteriophage SPP1]
1331	Bt1Gc1667	Bt1G3087	1415-2365	g2289093	1583	1486	2.60E-152	100	100	cell division protein tms26 - Bacillus subtilis [Bacillus subtilis]
1331	Bt1Gc1667	Bt1G3088	2441-3004	g586021	704	710	4.40E-70	66	99	(U76387) PRPP synthetase [Corynebacterium ammoniagenes]
1331	Bt1Gc1667	Bt1G3089	3072-3293	g586882	180	197	1.00E-15	47	99	PROBABLE PEPTIDYL-TRNA HYDROLASE (PTH) (STAGE V SPORULATION PROTEIN C) [Bacillus subtilis]
1331	Bt1Gc1667	Bt1G3090	3402-6932	g585481	4146	4003	0	68	100	HYPOTHETICAL 8.8 KD PROTEIN IN SPOVC-MFD INTERGENIC REGION [Bacillus subtilis]
1331	Bt1Gc1667	Bt1G3091	7069-7602	g586883	727	727	7.00E-72	77	99	TRANSCRIPTION-REPAIR COUPLING FACTOR (TRCF) [Bacillus subtilis]
1331	Bt1Gc1667	Bt1G3092	7851-9440	g4090864	745	624	5.70E-61	36	100	STAGE V SPORULATION PROTEIN T [Bacillus subtilis]
1331	Bt1Gc1667	Bt1G3093	9450-10287	g586885	912	917	5.10E-92	61	57	(AF023181) low temperature requirement B protein [Listeria monocytogenes]
1332	Bt1Gc1668	Bt1G3094	1270-568	g4126668	491	512	4.20E-49	42	100	HYPOTHETICAL 56.1 KD PROTEIN IN MFD-DIVIC INTERGENIC REGION [Bacillus subtilis]
1332	Bt1Gc1668	Bt1G3095	1988-1275	g1170075	434	497	1.60E-47	42	100	(AB016427) thioesterase II-like protein [Bacillus licheniformis]
1332	Bt1Gc1668	Bt1G3096	17853-1	g2623773	6940	7957	0	39	67	GSP PROTEIN [Brevibacillus brevis]
1333	Bt1Gc1679	Bt1G3097	1-215	g2634223	257	259	2.70E-22	68	93	(AF004835) tyrocidine synthetase 3 [Brevibacillus brevis]
1333	Bt1Gc1679	Bt1G3098	55-3110	g2815495	261	371	3.70E-34	35	67	(Z99113) yoeD [Bacillus subtilis]
1333	Bt1Gc1679	Bt1G3099	1595-2260	g1929337	468	477	2.20E-45	40	100	SPORE COAT PROTEIN B [Bacillus subtilis]
1333	Bt1Gc1679	Bt1G3100	4439-5156	g2633724	420	487	4.30E-46	38	33	(Z93767) ywrJ [Bacillus subtilis]
1333	Bt1Gc1679	Bt1G3101	6439-5945	g3261548	100	166	2.90E-12	26	61	(Z99111) similar to two-component sensor histidine kinase [Bacillus subtilis]
1333	Bt1Gc1679	Bt1G3102	7968-8747	g1805391	651	593	1.10E-57	48	100	(AL022003) hypothetical protein Rv1716 [Mycobacterium tuberculosis]
1333	Bt1Gc1679	Bt1G3103	10053-8453	g2352826	694	723	1.80E-71	35	100	(D50453) ycgL [Bacillus subtilis]
1333	Bt1Gc1679	Bt1G3104	11071-10212	g2226259	891	932	1.30E-93	56	57	(AF009224) benzoate transport protein [Acinetobacter sp. ADP1]
1333	Bt1Gc1679	Bt1G3105	11501-11204	g2255559	478	494	3.40E-47	99	21	(Y14084) hypothetical protein [Bacillus subtilis]
1334	Bt1Gc1682	Bt1G3106	767-1	g2226127	965	964	5.30E-97	71	81	ORF IS231C [Bacillus thuringiensis]
										(Y14078) Hypothetical protein [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1334	Bt1Gc1682	Bt1G3107	3742-833	g2226128	932	1121	1.20E-113	29	100	(Y14078) Hypothetical protein [Bacillus subtilis]
1334	Bt1Gc1682	Bt1G3108	3742-833	g2226128	932	1121	1.20E-113	29	100	(Y14078) Hypothetical protein [Bacillus subtilis]
1334	Bt1Gc1682	Bt1G3109	4974-3742	g2226129	801	819	1.20E-81	41	100	(Y14078) Hypothetical protein [Bacillus subtilis]
1334	Bt1Gc1682	Bt1G3110	5909-6043	g2443232	151	151	7.60E-11	62	98	(D86417) YfJ [Bacillus subtilis]
1334	Bt1Gc1682	Bt1G3111	7157-6174	g1731309	817	814	4.20E-81	47	100	HYPOTHETICAL 37.2 KD PROTEIN IN IDH-DEOR INTERGENIC REGION [Bacillus subtilis]
1334	Bt1Gc1682	Bt1G3112	7366-10511	g3242253	200	335	1.30E-28	26	14	(Z83857) ppsC [Mycobacterium tuberculosis]
1334	Bt1Gc1682	Bt1G3113	8427-10077	g2961083	255	329	1.00E-29	37	85	(AF033674) unknown [Pseudomonas marginalis pv. alfalfae]
1334	Bt1Gc1682	Bt1G3114	8433-9485	g731090	159	237	5.80E-20	25	100	UXU OPERON TRANSCRIPTIONAL REGULATOR [Escherichia coli]
1334	Bt1Gc1682	Bt1G3115	9196-10349	g2506999	386	369	6.00E-34	27	100	CYANATE TRANSPORT PROTEIN CYNX [synthetic construct]
1335	Bt1Gc1680	Bt1G3116	1-257	g223559	409	425	7.00E-40	99	18	ORF IS231C [Bacillus thuringiensis]
1335	Bt1Gc1680	Bt1G3117	950-3340	g3328629	140	228	5.30E-19	30	78	(AE001295) hypothetical protein [Chlamydia trachomatis]
1335	Bt1Gc1680	Bt1G3118	2938-2183	g1706579	134	155	1.70E-10	23	100	PROTEIN ECSC [Bacillus subtilis]
1335	Bt1Gc1680	Bt1G3119	5504-4320	g584719	1687	1604	8.10E-165	81	100	ACETATE KINASE (ACETOKINASE) [Bacillus subtilis]
1335	Bt1Gc1680	Bt1G3120	6793-5807	g3916031	856	879	5.40E-88	52	100	HYPOTHETICAL 37.4 KD PROTEIN IN ACKA-SSPA INTERGENIC REGION [Bacillus subtilis]
1335	Bt1Gc1680	Bt1G3121	7489-6989	g2293238	590	573	1.50E-55	68	99	(AF008220) YtgI [Bacillus subtilis]
1335	Bt1Gc1680	Bt1G3122	8022-7600	g2293237	370	284	6.10E-25	50	99	(AF008220) YtfJ [Bacillus subtilis]
1335	Bt1Gc1680	Bt1G3123	8876-9427	g2462088	857	821	7.60E-82	91	99	(Y11170) BC541A protein [Bacillus cereus]
1335	Bt1Gc1680	Bt1G3124	11482-9920	g2293233	1275	1262	1.40E-128	48	100	(AF008220) YtcJ [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3125	939-292	g2337808	858	849	8.20E-85	77	100	(Y13937) putative CfxE protein [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3126	1823-945	g2337807	995	1016	1.70E-102	64	100	(Y13937) YloQ protein [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3127	4043-2070	g2337806	1553	1550	4.30E-159	51	100	(Y13937) putative Pkn2 protein [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3128	4804-4046	g2337805	647	629	1.70E-61	49	100	(Y13937) putative Ptc1 protein [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3129	5900-4812	g3287951	1537	1537	1.00E-157	78	100	HYPOTHETICAL 41.6 KD PROTEIN IN FMT-SPOVM INTERGENIC REGION [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3130	7239-5905	g3915867	1385	1400	3.40E-143	59	100	SUN PROTEIN [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3131	8177-7227	g2337802	1145	1152	6.40E-117	69	100	(Y13937) putative Fmt protein [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3132	8674-8192	g2499922	451	484	3.90E-46	57	99	POLYPEPTIDE DEFORMYLASE (PDF) (FORMYLMETHIONINE DEFORMYLASE) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1336	Bt1Gc1678	Bt1G3133	11091-8689	g3183549	2513	2387	8.60E-248	59	100	PRIMOSOMAL PROTEIN N' (REPLICATION FACTOR Y) [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3134	12293-11085	g2337799	1279	1310	1.20E-133	63	100	(Y113937) putative Dfp protein [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3135	13418-12684	g2337797	812	827	1.80E-82	68	100	(Y113937) putative Gmk protein [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3136	14518-13646	g2337796	637	600	2.00E-58	45	100	(Y113937) YloC protein [Bacillus subtilis]
1336	Bt1Gc1678	Bt1G3137	16875-14669	g2337795	2618	2403	1.70E-249	70	83	(Y113937) putative PacL protein [Bacillus subtilis]
1337	Bt1Gc1681	Bt1G3138	2360-3145	g2497492	484	518	9.80E-50	40	100	URIDYLATE KINASE (UK) (URIDINE MONOPHOSPHATE KINASE) (UMP KINASE) [Synecocystis sp.]
1337	Bt1Gc1681	Bt1G3139	4411-3218	g1176955	693	697	1.10E-68	37	100	HYPOTHETICAL 45.8 KD PROTEIN IN ACDA-NARI INTERGENIC REGION [Bacillus subtilis]
1337	Bt1Gc1681	Bt1G3140	8755-7827	g1817536	649	681	5.20E-67	45	100	(D78508) YfiL [Bacillus subtilis]
1338	Bt1Gc1684	Bt1G3141	91-2455	g122819	1494	1525	1.90E-156	66	97	GLUTAMYL-TRNA REDUCTASE (GLUTR) [Bacillus subtilis]
1338	Bt1Gc1684	Bt1G3142	1452-2273	g3172392	828	784	6.30E-78	56	100	(AF064058) HemX [Bacillus stearothermophilus]
1338	Bt1Gc1684	Bt1G3143	2289-3230	g122839	1027	1041	3.70E-105	65	100	PORPHOBILINOGEN DEAMINASE (PBG) (HYDROXYMETHYLBILANE SYNTHASE) (HMBS) (PRE-UROPORPHYRINOGEN SYNTHASE) [Bacillus subtilis]
1338	Bt1Gc1684	Bt1G3144	3215-3988	g399875	248	221	2.90E-18	31	100	UROPORPHYRINOGEN-III SYNTHASE (UROS) (UROPORPHYRINOGEN-III COSYNTHETASE) (HYDROXYMETHYLBILANE HYDROLYASE (CYCLIZING)) [Bacillus subtilis]
1338	Bt1Gc1684	Bt1G3145	4000-4971	g399870	1279	1281	1.40E-130	75	100	DELTA-AMINOLEVULINIC ACID DEHYDRATASE (PORPHOBILINOGEN SYNTHASE) (ALAD) (ALADH) [Bacillus subtilis]
1338	Bt1Gc1684	Bt1G3146	4986-6275	g399784	1810	1775	6.10E-183	80	100	GLUTAMATE-1-SEMIALDEHYDE 2,1-AMINOMUTASE (GSA) (GLUTAMATE-1-SEMIALDEHYDE AMINOTRANSFERASE) (GSA-AT) [Bacillus subtilis]
1338	Bt1Gc1684	Bt1G3147	6378-7022	g421533	291	319	5.50E-30	36	37	normal spore coat assembly spoVID - Bacillus subtilis []
1338	Bt1Gc1684	Bt1G3148	6378-8447	g421533	286	319	5.50E-30	33	47	normal spore coat assembly spoVID - Bacillus subtilis []
1338	Bt1Gc1684	Bt1G3149	7446-8478	g586918	486	547	8.30E-53	30	100	HYPOTHETICAL 40.6 KD PROTEIN IN SPOVID 3'REGION (ORF2) [Bacillus subtilis]
1338	Bt1Gc1684	Bt1G3150	8980-11622	g68538	3783	3739	0	78	100	valine-tRNA ligase (EC 6.1.1.9) - Bacillus stearothermophilus []



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1338	Bt1Gc1684	Bt1G3151	11719-13008	g585155	1057	1063	1.70E-107	49	100	FOLYLPOLYGLUTAMATE SYNTHASE (FOLYLPOLY-GAMMA-GLUTAMATE SYNTHETASE) (FPGS) [Bacillus subtilis]
1339	Bt1Gc1683	Bt1G3152	1-840	g3821837	145	271	1.50E-23	23	85	putative beta1,3-glucosyltransferase WaaV [Escherichia coli]
1339	Bt1Gc1683	Bt1G3153	2792-5287	g2633184	610	529	5.50E-49	28	100	(Z99108) yfhO [Bacillus subtilis]
1339	Bt1Gc1683	Bt1G3154	2792-5287	g2633184	610	529	5.50E-49	28	100	(Z99108) yfhO [Bacillus subtilis]
1339	Bt1Gc1683	Bt1G3155	5402-6829	g2498262	970	890	3.70E-89	55	100	CSBB PROTEIN [Bacillus subtilis]
1339	Bt1Gc1683	Bt1G3156	6506-8056	g2443256	1535	1428	3.60E-146	56	100	(D86417) YfmM [Bacillus subtilis]
1339	Bt1Gc1683	Bt1G3157	8986-8114	g2635800	490	480	1.00E-45	38	100	(Z99120) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1339	Bt1Gc1683	Bt1G3158	9088-12252	g2632985	3348	2412	3.20E-256	63	100	(Z99107) similar to acriflavin resistance protein [Bacillus subtilis]
1339	Bt1Gc1683	Bt1G3159	13116-12280	g1351814	554	639	1.50E-62	44	78	HYPOTHETICAL PROTEIN HI1455 [Haemophilus influenzae Rd]
1340	Bt1Gc1688	Bt1G3160	165-2224	g141450	74	136	1.40E-08	32	29	HYPOTHETICAL 37.1 KD PROTEIN IN TRANSPOSON TN4556 [Transposon Tn4556]
1340	Bt1Gc1688	Bt1G3161	2590-1298	g2497382	2231	2231	2.90E-231	100	100	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1341	Bt1Gc1686	Bt1G3162	1-519	g3123253	611	629	1.70E-61	67	32	PHOSPHOENOLPYRUVATE CARBOXYKINASE [Bacillus subtilis]
1341	Bt1Gc1686	Bt1G3163	2679-1759	g1881343	209	191	1.20E-14	24	100	(AB001488) FUNCTION UNKNOWN, SIMILAR PRODUCT IN E. COLI. [Bacillus subtilis]
1341	Bt1Gc1686	Bt1G3164	3300-2983	g1731088	148	165	2.50E-12	31	99	HYPOTHETICAL 11.8 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1341	Bt1Gc1686	Bt1G3165	5258-4623	g2618860	169	225	1.10E-18	25	100	(AF017113) HPp(Ser-P) phosphatase [Bacillus subtilis]
1341	Bt1Gc1686	Bt1G3166	6019-5246	g1763710	502	509	8.80E-49	41	100	(Z83337) highly similar to phosphotransferase system regulator [Bacillus subtilis]
1341	Bt1Gc1686	Bt1G3167	6116-6766	g2618860	291	298	2.00E-26	33	100	(AF017113) HPp(Ser-P) phosphatase [Bacillus subtilis]
1341	Bt1Gc1686	Bt1G3168	14867-7457	g2633724	354	429	1.20E-39	40	31	(Z99111) similar to two-component sensor histidine kinase [Bacillus subtilis]
1341	Bt1Gc1686	Bt1G3169	10189-11001	g2145391	584	601	1.60E-58	41	100	(Y09476) YisY [Bacillus subtilis]
1341	Bt1Gc1686	Bt1G3170	12969-11188	g1658371	711	822	6.00E-82	33	100	(U67998) unknown [Sinorhizobium meliloti]
1342	Bt1Gc1689	Bt1G3171	1024-668	g2635223	415	416	6.30E-39	64	99	(Z99118) similar to hypothetical proteins [Bacillus subtilis]
1342	Bt1Gc1689	Bt1G3172	3305-1125	g2739435	2972	2946	5.00E-307	77	100	(U86377) (p)ppGpp synthetase [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1342	Bt1Gc1689	Bt1G3173	4028-3519	g2580524	718	663	4.20E-65	82	99	(U86377) adenine phosphoribosyltransferase; Apt [Bacillus subtilis]
1342	Bt1Gc1689	Bt1G3174	6415-4067	g2635226	2077	2013	3.70E-208	52	100	(Z99118) similar to single-strand DNA-specific exonuclease [Bacillus subtilis]
1342	Bt1Gc1689	Bt1G3175	7407-6520	g1881354	571	490	9.10E-47	37	100	(AB001488) FUNCTION UNKNOWN, SIMILAR PRODUCT IN SYNECHOCYSTIS AND E. COLI. [Bacillus subtilis]
1342	Bt1Gc1689	Bt1G3176	9825-7606	g3220156	2085	1942	1.20E-200	60	100	(AF024506) SecDF protein [Bacillus subtilis]
1342	Bt1Gc1689	Bt1G3177	10570-12123	g267023	1492	1348	1.10E-137	57	100	STAGE V SPORULATION PROTEIN B [Bacillus subtilis]
1342	Bt1Gc1689	Bt1G3178	12760-12204	g2635233	457	480	1.00E-45	48	84	(Z99118) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1343	Bt1Gc1687	Bt1G3179	1915-2685	g1177019	893	893	1.80E-89	66	100	HYPOTHETICAL OXIDOREDUCTASE IN PEPT-KATB INTERGENIC REGION [Bacillus subtilis]
1343	Bt1Gc1687	Bt1G3180	10573-12003	g2494077	1323	1317	2.10E-134	55	100	NADP-DEPENDENT GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (NON-PHOSPHORYLATING GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE) (Bacillus subtilis)
1343	Bt1Gc1687	Bt1G3181	12623-13728	g479960	369	497	1.60E-47	29	56	NADH oxidase - Thermoanaerobacter brockii []
1343	Bt1Gc1687	Bt1G3182	15053-14131	g3183453	409	457	2.80E-43	59	43	HYPOTHETICAL 38.9 KD PROTEIN IN DAT-SPO0E INTERGENIC REGION [Bacillus subtilis]
1343	Bt1Gc1687	Bt1G3183	14644-15053	g225559	646	661	6.90E-65	96	28	ORF IS231C [Bacillus thuringiensis]
1344	Bt1Gc1685	Bt1G3184	1531-2589	g2829831	1281	1225	1.20E-124	68	100	D-ALANINE--D-ALANINE LIGASE (D-ALANYLALANINE SYNTHETASE) [Bacillus subtilis]
1344	Bt1Gc1685	Bt1G3185	2655-4043	g3024202	1276	1321	7.90E-135	56	100	UDP-N-ACETYLMURAMOYLALANYL-D-GLUTAMYL-2,6-DIAMINOPIMELATE--D-ALANYL-D-ALANYL LIGASE (UDP-MURNAC-PENTAPEPTIDE SYNTHETASE) (D-ALANYL-D-ALANINE-ADDING ENZYME) [Bacillus subtilis]
1344	Bt1Gc1685	Bt1G3186	3971-5806	g1881268	1765	1616	4.30E-166	70	100	(AB001488) ATP-DEPENDENT RNA HELICASE DEAD HOMOLOG. [Bacillus subtilis]
1344	Bt1Gc1685	Bt1G3187	6009-6971	g1176954	548	548	6.50E-53	37	100	HYPOTHETICAL 36.9 KD PROTEIN IN ACDA 5'REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1344	Bt1Gc1685	Bt1G3188	8412-6973	g1881271	249	214	1.60E-17	32	99	(AB001488) FUNCTION UNKNOWN, SIMILAR PRODUCT IN H. INFLUENZAE AND SYNECHOCYSTIS. [Bacillus subtilis]
1344	Bt1Gc1685	Bt1G3189	7630-7986	g1881272	304	285	4.80E-25	54	99	(AB001488) FUNCTION UNKNOWN, SIMILAR PRODUCT IN E. COLI AND MYCOPLASMA PNEUMONIAE. [Bacillus subtilis]
1344	Bt1Gc1685	Bt1G3190	7971-9093	g1881273	975	925	7.30E-93	52	100	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
1344	Bt1Gc1685	Bt1G3191	9220-10392	g2506170	1115	1141	9.40E-116	56	100	ALANINE RACEMASE []
1344	Bt1Gc1685	Bt1G3192	10692-10976	g1881275	227	209	5.40E-17	54	99	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
1344	Bt1Gc1685	Bt1G3193	10984-11331	g1881276	540	540	4.60E-52	93	99	(AB001488) FUNCTION UNKNOWN, SIMILAR PRODUCT IN MYCOBACTERIUM TUBERCULOSIS AND STAPHYLOCOCCUS AUREUS. [Bacillus subtilis]
1344	Bt1Gc1685	Bt1G3194	11411-12719	g2632778	1293	1322	6.20E-135	59	56	(Z99106) similar to hypothetical proteins [Bacillus subtilis]
1345	Bt1Gc1690	Bt1G3195	697-1	g1881335	893	715	1.30E-70	78	80	(AB001488) SIMILAR TO YQFU, YXKD, YITB OF B. SUBTILIS. [Bacillus subtilis]
1345	Bt1Gc1690	Bt1G3196	4017-3001	g3928722	504	478	1.70E-45	36	100	(AL034355) putative cytochrome oxidase subunit II [Streptomyces coelicolor]
1345	Bt1Gc1690	Bt1G3197	5362-3982	g2829796	999	976	2.90E-98	44	100	CYTOCHROME D UBIQUINOL OXIDASE SUBUNIT I [Bacillus subtilis]
1345	Bt1Gc1690	Bt1G3198	5590-6150	g2293156	693	611	1.40E-59	66	99	(AF008220) YtIB [Bacillus subtilis]
1345	Bt1Gc1690	Bt1G3199	6386-6162	g2293294	327	329	1.00E-29	75	99	(AF008220) YtjA [Bacillus subtilis]
1345	Bt1Gc1690	Bt1G3200	6515-6985	g2293157	702	702	3.10E-69	82	99	(AF008220) YtjB [Bacillus subtilis]
1345	Bt1Gc1690	Bt1G3201	8745-9218	g2293161	514	517	1.20E-49	61	99	(AF008220) YtkD [Bacillus subtilis]
1345	Bt1Gc1690	Bt1G3202	12653-11844	g2293295	858	639	1.50E-62	62	100	(AF008220) YtID [Bacillus subtilis]
1345	Bt1Gc1690	Bt1G3203	13444-12646	g2492576	453	494	3.40E-47	40	100	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN MJ0412 [Methanococcus jannaschii]
1345	Bt1Gc1690	Bt1G3204	13724-13410	g2293297	292	296	3.30E-26	52	99	(AF008220) YtIB. [Bacillus subtilis]
1345	Bt1Gc1690	Bt1G3205	14417-13724	g2293298	698	636	3.10E-62	62	100	(AF008220) putative transcription regulator [Bacillus subtilis]
1346	Bt1Gc1691	Bt1G3206	377-1	g3913706	522	483	5.00E-46	79	86	FERRIC UPTAKE REGULATION PROTEIN HOMOLOG 3 [Bacillus subtilis]
1346	Bt1Gc1691	Bt1G3207	1133-663	g2983147	385	408	4.40E-38	49	99	(AE000692) hypothetical protein [Aquifex aeolicus]
1346	Bt1Gc1691	Bt1G3208	4374-5660	g2829488	1802	1731	2.80E-178	80	100	GLUTAMATE-1-SEMIALDEHYDE 2,1-AMINOMUTASE 2 (GSA) (GLUTAMATE-1-SEMIALDEHYDE AMINOTRANSFERASE) (GSA-AT) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1346	Bt1Gc1691	Bt1G3210	5322-10311	g4521156	812	657	1.80E-125	29	100	(AB024708) glutamine 2-oxoglutarate aminotransferase large subunit [Corynebacterium glutamicum]
1346	Bt1Gc1691	Bt1G3211	5916-10331	g2983926	874	908	7.10E-129	29	100	(AE000746) glutamate synthase large subunit [Aquifex aeolicus]
1346	Bt1Gc1691	Bt1G3212	5322-10311	g4521156	812	657	1.80E-125	29	100	(AB024708) glutamine 2-oxoglutarate aminotransferase large subunit [Corynebacterium glutamicum]
1347	Bt1Gc1693	Bt1G3213	2027-366	g3121769	1433	1331	6.90E-136	53	100	POTASSIUM-TRANSPORTING ATPASE A CHAIN (ATP PHOSPHOHYDROLASE [Clostridium acetobutylicum]
1347	Bt1Gc1693	Bt1G3214	4506-3691	g2633525	788	617	3.20E-60	56	100	(Z99110) similar to phosphomethylpyrimidine kinase [Bacillus subtilis]
1347	Bt1Gc1693	Bt1G3215	5541-4525	g2633524	985	1024	2.30E-103	58	100	(Z99110) similar to thiamin biosynthesis [Bacillus subtilis]
1347	Bt1Gc1693	Bt1G3216	6310-5543	g2633523	1027	1027	1.10E-103	80	100	(Z99110) similar to thiamin biosynthesis [Bacillus subtilis]
1347	Bt1Gc1693	Bt1G3217	7635-6510	g2633521	524	520	6.00E-50	33	100	(Z99110) similar to sarcosine oxidase [Bacillus subtilis]
1347	Bt1Gc1693	Bt1G3218	8242-7628	g135587	446	456	3.60E-43	43	100	REGULATORY PROTEIN TENI [Bacillus subtilis]
1347	Bt1Gc1693	Bt1G3219	10003-7522	g4377342	434	485	3.10E-46	26	100	(AE001681) ABC transporter permease [Chlamydia pneumoniae]
1347	Bt1Gc1693	Bt1G3220	11440-10027	g2492576	391	446	4.20E-42	39	100	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN MJ0412 [Methanococcus jannaschii]
1347	Bt1Gc1693	Bt1G3221	11422-10772	g2314455	562	469	1.50E-44	47	100	(AE000633) transcriptional regulator (tenA) [Helicobacter pylori 26695]
1347	Bt1Gc1693	Bt1G3222	12863-13192	g2984049	142	227	6.70E-19	44	52	(AE000754) lipoprotein NlpD fragment [Aquifex aeolicus]
1347	Bt1Gc1693	Bt1G3223	15684-15241	g2339998	350	312	6.60E-28	49	99	(Z98682) YlbB protein [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3224	1985-489	g586058	2144	2023	3.20E-209	80	100	LYSYL-TRNA SYNTHETASE (LYSINE--TRNA LIGASE) (LYSRS) [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3225	3137-2142	g586896	1395	1404	1.30E-143	81	100	HYPOTHETICAL 37.1 KD PROTEIN IN FOLK-LYSS INTERGENIC REGION [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3226	3831-3319	g585263	516	503	3.80E-48	59	99	2-AMINO-4-HYDROXY-6-HYDROXYMETHYLDIHYDROPTERIDINE PYROPHOSPHOKINASE (7,8-DIHYDRO-6-HYDROXYMETHYLPTERIN-PYROPHOSPHOKINASE) (HPPK) (6-HYDROXYMETHYL-7,8-DIHYDROPTERIN PYROPHOSPHOKINASE) (ppPK) [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3227	4190-3831	g141435	439	396	8.30E-37	62	99	DIHYDRONEOPTERIN ALDOLASE (DHNA) [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3228	5039-4186	g2126927	903	839	9.40E-84	62	100	dihydropteroate pyrophosphorylase - Bacillus subtilis

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1348	Bt1Gc1694	Bt1G3229	5895-5024	g129533	710	747	5.30E-74	49	100	[Bacillus subtilis] 4-AMINO-4-DEOXYCHORISMATE LYASE (ADC LYASE) [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3230	6479-5898	g129521	789	793	7.10E-79	77	99	PARA-AMINOBENZOATE SYNTHASE GLUTAMINE AMIDOTRANSFERASE COMPONENT II (ADC SYNTHASE) / ANTHRANILATE SYNTHASE COMPONENT II [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3231	7882-6482	g129527	1384	1232	2.10E-125	58	100	PARA-AMINOBENZOATE SYNTHASE COMPONENT I (ADC SYNTHASE) [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3232	9031-8111	g585031	1268	1132	8.40E-115	81	100	CYSTEINE SYNTHASE (O-ACETYL SERINE SULFHYDRYLASE) (O-ACETYL SERINE (THIOL)-LYASE) (CSASE) (SUPEROXIDE-INDUCIBLE PROTEIN 11) (SOI11) [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3233	10014-9137	g586894	923	906	7.50E-91	62	100	HYPOTHETICAL 31.8 KD PROTEIN IN FTSH-CYSK INTERGENIC REGION [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3234	10812-10114	g586893	892	898	5.30E-90	70	100	HYPOTHETICAL 26.2 KD PROTEIN IN FTSH-CYSK INTERGENIC REGION [Bacillus subtilis]
1348	Bt1Gc1694	Bt1G3235	12007-11045	g585159	1303	1234	1.30E-125	79	50	CELL DIVISION PROTEIN FTSH HOMOLOG [Bacillus subtilis]
1349	Bt1Gc1697	Bt1G3236	1-865	g2226185	489	510	6.90E-49	36	86	(Y14081) hypothetical protein [Bacillus subtilis]
1349	Bt1Gc1697	Bt1G3237	1241-1972	g2226242	756	761	1.70E-75	59	100	(Y14083) hypothetical protein [Bacillus subtilis]
1349	Bt1Gc1697	Bt1G3238	1985-2971	g2226243	1264	1274	7.60E-130	72	100	(Y14083) hypothetical protein [Bacillus subtilis]
1349	Bt1Gc1697	Bt1G3239	3866-5392	g2226245	1559	1547	8.90E-159	58	100	(Y14083) hypothetical protein [Bacillus subtilis]
1349	Bt1Gc1697	Bt1G3240	5982-6968	g4323056	1187	1122	9.70E-114	68	100	(AF098509) putative sulfur-binding protein [Enterobacter cloacae]
1349	Bt1Gc1697	Bt1G3241	6997-7827	g118163	733	580	2.60E-56	51	100	SULFATE TRANSPORT SYSTEM PERMEASE PROTEIN CYST [Escherichia coli]
1349	Bt1Gc1697	Bt1G3242	7857-8732	g118165	725	707	9.20E-70	48	100	SULFATE TRANSPORT SYSTEM PERMEASE PROTEIN CYSW [Escherichia coli]
1349	Bt1Gc1697	Bt1G3243	9088-9786	g2462099	1142	1081	2.10E-109	95	100	(Y10908) sulfate transport ATP-binding protein [Bacillus cereus]
1350	Bt1Gc1696	Bt1G3244	330-1	g2633746	496	496	2.10E-47	82	66	(Z99111) similar to hypothetical proteins [Bacillus subtilis]
1350	Bt1Gc1696	Bt1G3245	1068-352	g2633745	818	833	4.10E-83	64	100	(Z99111) similar to coenzyme PQQ synthesis [Bacillus subtilis]
1350	Bt1Gc1696	Bt1G3246	1459-1079	g1175542	147	202	3.00E-16	35	99	HYPOTHETICAL PROTEIN HI1190 [Haemophilus

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrq	NCBI gi description
1350	Bt1Gc1696	Bt1G3247	2211-1555	g2633743	862	871	3.80E-87	71	100	influenzae Rd] (Z99111) similar to hypothetical proteins [Bacillus subtilis]
1350	Bt1Gc1696	Bt1G3248	6361-4667	g323354	1325	1257	5.20E-138	45	100	(AE001270) oligoendopeptidase F, putative [Treponema pallidum]
1350	Bt1Gc1696	Bt1G3249	6937-6503	g2633738	381	358	8.80E-33	54	99	(Z99111) similar to transcriptional regulator (MarR family) [Bacillus subtilis]
1350	Bt1Gc1696	Bt1G3250	7130-8650	g2633737	914	966	3.30E-97	37	100	(Z99111) similar to two-component sensor histidine kinase [Bacillus subtilis]
1350	Bt1Gc1696	Bt1G3251	9441-10271	g1176306	264	344	2.70E-31	30	100	HYPOTHETICAL PROTEIN HI0755 []
1350	Bt1Gc1696	Bt1G3252	11764-11144	g2633732	372	400	3.10E-37	39	100	(Z99111) similar to hypothetical proteins [Bacillus subtilis]
1350	Bt1Gc1696	Bt1G3253	13141-9996	g4894270	335	437	3.70E-41	39	57	(AF065404) pXO1-54 [Bacillus anthracis]
1350	Bt1Gc1696	Bt1G3254	15206-12892	g3123226	421	466	1.60E-43	36	39	INTERNALIN A PRECURSOR []
1350	Bt1Gc1696	Bt1G3255	15593-14213	g2633809	238	249	3.10E-21	29	100	(Z99111) transcriptional regulator (DeoR family) [Bacillus subtilis]
1350	Bt1Gc1696	Bt1G3256	17461-16819	g1763079	596	642	7.10E-63	59	58	(U69493) PhnW [Salmonella typhimurium]
1351	Bt1Gc1698	Bt1G3257	1-1444	g2634068	2120	1851	5.40E-191	88	93	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
1351	Bt1Gc1698	Bt1G3258	1610-2401	g2634069	1021	1021	4.90E-103	72	100	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
1351	Bt1Gc1698	Bt1G3259	2554-2811	g1174408	382	312	6.60E-28	92	99	STAGE V SPOULATION PROTEIN S [Bacillus subtilis]
1351	Bt1Gc1698	Bt1G3260	3023-3793	g2062103	583	609	2.20E-59	47	100	(X93169) orfD [Listeria monocytogenes]
1351	Bt1Gc1698	Bt1G3261	4027-5787	g2129427	597	709	1.60E-83	34	100	2-oxoacid ferredoxin oxidoreductase (EC 1.-.-.-) alpha chain - Sulfolobus acidocaldarius [Sulfolobus sp. 7]
1351	Bt1Gc1698	Bt1G3262	6269-6628	g2462119	609	609	2.20E-59	98	99	(Y11140) bc48a [Bacillus cereus]
1351	Bt1Gc1698	Bt1G3263	7062-8588	g2634073	2124	2124	6.40E-220	78	100	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
1351	Bt1Gc1698	Bt1G3264	8595-9023	g2634074	484	485	3.10E-46	66	99	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
1351	Bt1Gc1698	Bt1G3265	9153-9692	g116957	562	500	7.90E-48	58	99	SPORE COAT PROTEIN E [Bacillus subtilis]
1351	Bt1Gc1698	Bt1G3266	9876-12539	g2634076	2799	2761	2.00E-287	65	100	(Z99112) mutS [Bacillus subtilis]
1351	Bt1Gc1698	Bt1G3267	12557-14497	g1709188	1927	1248	8.60E-188	60	100	DNA MISMATCH REPAIR PROTEIN MUTL [Bacillus subtilis]
1351	Bt1Gc1698	Bt1G3268	20161-18279	g1731026	1650	1685	2.10E-173	51	98	HYPOTHETICAL 73.2 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
1352	Bt1Gc1699	Bt1G3269	324-1346	g2635771	1170	1136	3.20E-115	66	100	(Z99120) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis]
1352	Bt1Gc1699	Bt1G3270	1339-2004	g2635770	773	479	1.30E-45	68	100	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1352	Bt1Gc1699	Bt1G3271	2030-3715	g2635769	793	787	3.10E-78	62	100	(Z99120) similar to hypothetical proteins [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1352	Bt1Gc1699	Bt1G3272	3957-4739	g3183561	1221	1221	3.10E-124	90	100	VEGETATIVE PROTEIN 296 (VEG296) [Bacillus subtilis]
1352	Bt1Gc1699	Bt1G3273	4758-6047	g2635766	1445	1339	9.80E-137	67	100	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1352	Bt1Gc1699	Bt1G3274	6050-7267	g2635765	1726	1674	3.10E-172	78	100	(Z99120) similar to NifS protein homolog [Bacillus subtilis]
1352	Bt1Gc1699	Bt1G3275	7740-9134	g2635763	2232	2136	3.40E-221	91	100	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1352	Bt1Gc1699	Bt1G3276	9816-11195	g2635734	1027	1019	7.90E-103	45	100	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1352	Bt1Gc1699	Bt1G3277	11285-11587	g2635733	345	350	6.20E-32	70	99	(Z99120) yunC [Bacillus subtilis]
1353	Bt1Gc1700	Bt1G3278	1-214	g136144	316	340	7.10E-31	94	15	TRANSPORPOSE FOR INSERTION SEQUENCE ELEMENT IS231C [Bacillus thuringiensis]
1353	Bt1Gc1700	Bt1G3279	4477-2081	g1944414	2976	2926	6.60E-305	71	100	(D87026) glycogen phosphorylase [Bacillus stearothermophilus]
1353	Bt1Gc1700	Bt1G3280	5926-4478	g2811062	1589	1608	3.10E-165	62	100	GLYCOGEN SYNTHASE (STARCH (BACTERIAL GLYCOGEN) SYNTHASE) [Bacillus stearothermophilus]
1353	Bt1Gc1700	Bt1G3281	7075-6050	g2811061	992	923	1.20E-92	57	100	GLYCOGEN BIOSYNTHESIS PROTEIN GLGD [Bacillus stearothermophilus]
1353	Bt1Gc1700	Bt1G3282	8218-7079	g729582	1475	1484	4.20E-152	71	100	GLUCOSE-1-PHOSPHATE ADENYL-TRANSFERASE (ADP-GLUCOSE SYNTHASE) (ADP-GLUCOSE PYROPHOSPHORYLASE) [Bacillus subtilis]
1353	Bt1Gc1700	Bt1G3283	10106-8187	g98191	2370	2269	2.80E-235	65	100	1,4-alpha-glucan branching enzyme (EC 2.4.1.18) - Bacillus stearothermophilus [Bacillus stearothermophilus]
1353	Bt1Gc1700	Bt1G3284	12485-11529	g126054	1137	1109	2.30E-112	66	100	L-LACTATE DEHYDROGENASE X []
1354	Bt1Gc1692	Bt1G3285	1640-260	g282286	427	514	2.60E-49	34	100	VanS - Enterococcus faecium []
1354	Bt1Gc1692	Bt1G3286	2067-1354	g4481749	539	547	8.30E-53	45	100	(AF007865) BacR [Bacillus licheniformis]
1354	Bt1Gc1692	Bt1G3287	3432-2323	g2293150	1034	1051	3.20E-106	54	100	(AF008220) Ytd [Bacillus subtilis]
1354	Bt1Gc1692	Bt1G3288	3965-3600	g1149666	367	371	3.70E-34	60	99	(X86498) IS1136 DNA [Clostridium perfringens]
1354	Bt1Gc1692	Bt1G3289	4747-2525	g3426013	330	510	6.90E-49	30	94	(AB016803) transposase [Deinococcus radiodurans]
1354	Bt1Gc1692	Bt1G3290	4741-3576	g1002992	318	460	1.40E-43	34	100	(U34772) ORF375 [Dichelobacter nodosus]
1354	Bt1Gc1692	Bt1G3291	5349-4059	g2495742	390	418	3.90E-39	40	100	HYPOTHETICAL PROTEIN MJ0014 [Methanococcus jannaschii]
1354	Bt1Gc1692	Bt1G3292	6837-5386	g2293149	1263	1197	1.10E-121	51	100	(AF008220) OSB-CoA synthase [Bacillus subtilis]
1354	Bt1Gc1692	Bt1G3293	7817-7005	g2293148	1233	1118	2.60E-113	84	100	(AF008220) dihydroxynaphthoate synthase [Bacillus subtilis]
1354	Bt1Gc1692	Bt1G3294	8721-7896	g2293147	568	565	1.00E-54	43	100	(AF008220) YtxM [Bacillus subtilis]
1354	Bt1Gc1692	Bt1G3295	10426-8711	g2293146	1355	1305	3.90E-133	46	99	(AF008220) SHCHC synthase [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3296	1376-2221	g2635736	932	781	1.30E-77	56	100	(Z99120) yunF [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1355	Bt1Gc1695	Bt1G3297	5772-3403	g2500765	301	395	1.10E-36	29	67	SENSOR PROTEIN VANSB (VANCOMYCIN B-TYPE RESISTANCE PROTEIN VANSB) [VANCOMYCIN HISTIDINE PROTEIN KINASE] [Enterococcus faecalis]
1355	Bt1Gc1695	Bt1G3298	6227-5523	g2126871	923	941	1.50E-94	78	100	purine nucleoside phosphorylase (EC 2.7.-.-) II - Bacillus stearothermophilus [Bacillus stearothermophilus]
1355	Bt1Gc1695	Bt1G3299	9033-7195	g2804528	1551	1492	6.00E-153	51	100	(D85082) YfiX [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3300	9777-9019	g2633166	430	415	8.00E-39	38	100	(Z99108) yfiW [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3301	10320-10859	g466187	332	366	1.20E-33	41	99	HYPOTHETICAL 20.4 KD PROTEIN IN RIBT-DACB INTERGENIC REGION (ORFX9) [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3302	10983-11882	g2619018	762	752	1.60E-74	52	100	(AF027868) superoxide dismutase [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3303	11989-13124	g461914	851	796	3.40E-79	48	100	PENICILLIN-BINDING PROTEIN 5* PRECURSOR (D-ALANYL-D-ALANINE CARBOXYPEPTIDASE) (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) (PBP-5*) [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3304	13115-13702	g466188	705	654	3.80E-64	67	99	SPORE MATURATION PROTEIN A [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3305	13708-14241	g466189	578	614	6.60E-60	66	99	SPORE MATURATION PROTEIN B [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3306	14528-15214	g466190	826	831	6.60E-83	69	100	RIBOSOMAL LARGE SUBUNIT PSEUDOURIDINE SYNTHASE B (PSEUDOURIDYLATE SYNTHASE) (URACIL HYDROLYASE) [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3307	15353-15896	g466191	427	462	8.40E-44	46	99	RESA PROTEIN [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3308	16009-17641	g466192	1836	1807	2.50E-186	63	100	RESB PROTEIN [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3309	17744-18800	g466193	1286	1293	7.30E-132	68	100	RESC PROTEIN [Bacillus subtilis]
1355	Bt1Gc1695	Bt1G3310	19101-19820	g466194	965	844	2.80E-84	77	100	TRANSCRIPTIONAL REGULATORY PROTEIN RESD [Bacillus subtilis]
1356	Bt1Gc1702	Bt1G3311	438-1631	g2636487	775	675	2.30E-66	42	100	(Z99124) pyrimidine-nucleoside transport protein [Bacillus subtilis]
1356	Bt1Gc1702	Bt1G3312	4121-5350	g2507254	1410	1415	8.60E-145	64	100	AMINOPEPTIDASE AMPS [Bacillus subtilis]
1356	Bt1Gc1702	Bt1G3313	7838-6396	g120777	1413	1330	8.80E-136	56	100	SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+) (SSDH) [Escherichia coli]
1356	Bt1Gc1702	Bt1G3314	9231-7852	g1072970	797	788	2.40E-78	39	100	dhIR protein - Xanthobacter autotrophicus [Xanthobacter autotrophicus]
1356	Bt1Gc1702	Bt1G3315	10660-9353	g1805459	1472	1374	1.90E-140	63	100	(D50453) homologue of 4-aminobutyrate aminotransferase GabT of E. coli [Bacillus subtilis]
1356	Bt1Gc1702	Bt1G3316	12624-11719	g2632986	1023	1038	7.70E-105	65	100	(Z99107) similar to hypothetical proteins [Bacillus subtilis]
1356	Bt1Gc1702	Bt1G3317	14612-13188	g2589196	2081	2090	2.60E-216	86	100	(AF008553) Glu-tRNA <sup>Gln</sup> amidotransferase subunit B



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1356	Bt1Gc1702	Bt1G3318	16084-14630	g2632982	2025	1875	1.60E-193	80	100	[Bacillus subtilis] [Z99107] alternate gene name: yedB; similar to amidase [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3319	1680-1026	g322156	384	388	5.80E-36	42	100	hypothetical protein 1 - Bacillus megaterium [Bacillus megaterium]
1357	Bt1Gc1701	Bt1G3320	3329-1881	g1881270	530	655	3.00E-64	32	100	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3321	3792-3316	g1881269	359	363	2.60E-33	43	99	(AB001488) FUNCTION UNKNOWN. [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3322	5664-4954	g1706579	594	497	1.60E-47	49	100	PROTEIN ECSC [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3323	6895-5169	g1706578	609	610	1.70E-59	36	100	PROTEIN ECSB [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3324	7625-6885	g1706577	892	851	5.00E-85	68	100	ABC-TYPE TRANSPORTER ATP-BINDING PROTEIN ECSA [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3325	8194-8628	g3023940	524	491	7.10E-47	66	99	HIT PROTEIN [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3326	8805-9158	g2226119	214	215	1.30E-17	31	99	(Y14077) Hypothetical protein [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3327	9467-11001	g132549	615	659	1.10E-64	60	100	PROTEASE PRODUCTION REGULATORY PROTEIN HPR [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3328	10368-10030	g2226121	210	226	8.60E-19	36	99	(Y14077) Hypothetical protein [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3329	11096-11980	g131027	564	284	5.00E-33	45	100	PROTEIN EXPORT PROTEIN PRSA PRECURSOR [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3330	12396-14729	g2635858	1830	1833	4.40E-189	46	100	(Z99121) yvgS [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3331	16936-16415	g3024615	156	208	6.90E-17	25	99	RNA POLYMERASE SIGMA FACTOR SIGV [Bacillus subtilis]
1357	Bt1Gc1701	Bt1G3332	16972-16412	g548832	157	172	4.50E-13	26	99	RNA POLYMERASE SIGMA-H FACTOR (SIGMA-30) [Pseudomonas aeruginosa]
1358	Bt1Gc1704	Bt1G3333	438-1139	g586808	957	864	2.10E-86	79	100	HYPOTHETICAL 27.2 KD SENSORY TRANSDUCTION PROTEIN IN ROCR-PURA INTERGENIC REGION [Bacillus subtilis]
1358	Bt1Gc1704	Bt1G3334	1146-2990	g1064813	1694	1649	1.40E-169	55	100	(D78193) homologous to sp:PHOR_BACSU [Bacillus subtilis]
1358	Bt1Gc1704	Bt1G3335	2984-4291	g2636586	264	429	2.60E-40	26	100	(Z99124) yycH [Bacillus subtilis]
1358	Bt1Gc1704	Bt1G3336	5107-5910	g1064810	979	983	5.20E-99	71	100	(D78193) yycJ [Bacillus subtilis]
1358	Bt1Gc1704	Bt1G3337	5962-7146	g1731364	929	929	2.70E-93	48	100	HYPOTHETICAL PROTEASE IN ROCR-PURA INTERGENIC REGION [Bacillus subtilis]
1358	Bt1Gc1704	Bt1G3338	7439-7915	g1064795	640	640	1.20E-62	77	99	(D78193) yycA [Bacillus subtilis]
1358	Bt1Gc1704	Bt1G3339	6807-10716	g2633724	353	436	1.70E-43	36	35	(Z99111) similar to two-component sensor histidine kinase [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1358	Bt1Gc1704	Bt1G3340	10318-11100	g4155054	207	272	1.10E-23	29	100	(AE001484) putative [Helicobacter pylori J99]
1358	Bt1Gc1704	Bt1G3341	12896-11919	g1934831	1518	1447	3.50E-148	88	100	(Z93939) unknown [Bacillus subtilis]
1358	Bt1Gc1704	Bt1G3342	13379-14084	g1177010	376	428	3.40E-40	65	26	PROBABLE RNA HELICASE IN WAPA-LICT INTERGENIC REGION [Bacillus subtilis]
1358	Bt1Gc1704	Bt1G3343	13779-14084	g225559	467	488	1.50E-46	97	21	ORF IS231C [Bacillus thuringiensis]
1359	Bt1Gc1703	Bt1G3344	1450-98	g2226203	1319	1054	1.60E-106	56	100	(Y14082) hypothetical protein [Bacillus subtilis]
1359	Bt1Gc1703	Bt1G3345	2483-1998	g2462962	321	323	4.50E-29	35	99	(AJ000974) YlnF protein [Bacillus subtilis]
1359	Bt1Gc1703	Bt1G3346	3991-3278	g267052	987	900	3.20E-90	76	100	UROPORPHYRIN-III C-METHYLTRANSFERASE (UROGEN III METHYLASE) (SUMT) (UROPORPHYRINOGEN III METHYLASE) (UROM) [Bacillus megaterium]
1359	Bt1Gc1703	Bt1G3347	5903-4369	g730154	857	928	3.50E-93	42	100	FERREDOXIN--NITRITE REDUCTASE [Synechococcus sp.]
1359	Bt1Gc1703	Bt1G3348	7672-6524	g2462958	1364	1348	1.10E-137	68	100	(AJ000974) putative sulfate adenylyltransferase [Bacillus subtilis]
1359	Bt1Gc1703	Bt1G3349	8407-7709	g2462956	821	823	4.70E-82	59	100	(AJ000974) putative phospho-adenylylsulphate sulfotransferase [Bacillus subtilis]
1359	Bt1Gc1703	Bt1G3350	10252-8945	g1075836	1411	1426	5.90E-146	62	98	diaminopimelate decarboxylase (EC 4.1.1.20) - Bacillus subtilis [Bacillus subtilis]
1359	Bt1Gc1703	Bt1G3351	11354-13254	g1731026	1826	1797	2.90E-185	55	100	HYPOTHETICAL 73.2 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
1359	Bt1Gc1703	Bt1G3352	13420-14109	g2493570	300	330	8.20E-30	34	100	CDP-DIACYLGLYCEROL--SERINE O-PHOSPHATIDYLTRANSFERASE (PHOSPHATIDYL SERINE SYNTHASE) [Helicobacter pylori 26695]
1359	Bt1Gc1703	Bt1G3353	15751-14146	g3257003	732	742	1.80E-73	45	100	(AP000002) 376aa long hypothetical dehydrogenase [Pyrococcus horikoshii]
1359	Bt1Gc1703	Bt1G3354	17194-16575	g3913834	493	489	1.20E-46	50	100	PHOSPHORIBOSYL-AMP CYCLOHYDROLASE / PHOSPHORIBOSYL-ATP PYROPHOSPHOHYDROLASE [Bacillus subtilis]
1359	Bt1Gc1703	Bt1G3355	17936-17175	g3913862	812	836	2.00E-83	63	100	HISF PROTEIN (CYCLASE) [Aquifex aeolicus]
1359	Bt1Gc1703	Bt1G3356	18658-17933	g3913860	512	538	7.40E-52	43	100	PHOSPHORIBOSYLFORMIMINO-5-AMINOIMIDAZOLE CARBOXAMIDE RIBOTIDE ISOMERASE [Bacillus subtilis]
1359	Bt1Gc1703	Bt1G3357	19257-18655	g399894	485	501	6.20E-48	49	100	AMIDOTRANSFERASE HISH [Lactococcus lactis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat. nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1359	Bt1Gc1703	Bt1G3358	19836-19255	g2618867	495	495	2.70E-47	51	99	(AF017113) imidazoleglycerol-phosphate dehydratase [Bacillus subtilis]
1359	Bt1Gc1703	Bt1G3359	20555-19836	g3122219	666	644	4.30E-63	57	56	HISTIDINOL DEHYDROGENASE (HDH) [Bacillus subtilis]
1360	Bt1Gc1705	Bt1G3360	14-781	g2635726	907	868	8.00E-87	69	100	(Z99120) similar to N-acetyl-glucosamine catabolism [Bacillus subtilis]
1360	Bt1Gc1705	Bt1G3361	1631-1036	g1175888	343	368	7.70E-34	37	92	PUTATIVE NAD(P)H OXIDOREDUCTASE HI1544 [Haemophilus influenzae Rd]
1360	Bt1Gc1705	Bt1G3362	6376-4777	g1695869	180	289	1.80E-25	37	99	(U59239) orf3 [Serratia marcescens]
1360	Bt1Gc1705	Bt1G3363	7725-7228	g2635725	582	520	6.00E-50	67	99	(Z99120) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1360	Bt1Gc1705	Bt1G3364	7834-8850	g2635724	705	720	3.80E-71	41	100	(Z99120) yutH [Bacillus subtilis]
1360	Bt1Gc1705	Bt1G3365	9641-9313	g2635719	347	372	2.90E-34	68	99	(Z99120) similar to NifU protein homolog [Bacillus subtilis]
1360	Bt1Gc1705	Bt1G3366	9733-10062	g2635718	257	281	1.30E-24	47	99	(Z99120) yuzD [Bacillus subtilis]
1360	Bt1Gc1705	Bt1G3367	10124-10903	g4894249	297	372	2.90E-34	31	100	(AF065404) pXO1-33 [Bacillus anthracis]
1361	Bt1Gc1708	Bt1G3368	1-1031	g1176961	958	914	1.10E-91	54	90	HYPOTHETICAL 39.4 KD PROTEIN IN GNTR-HTPG INTERGENIC REGION [Bacillus subtilis]
1361	Bt1Gc1708	Bt1G3369	1245-2132	g1703387	1118	1139	1.50E-115	74	100	ARGINASE [Bacillus caldovelox]
1361	Bt1Gc1708	Bt1G3370	2384-3202	g1944009	978	978	1.80E-98	68	100	(AB002150) YbbP [Bacillus subtilis]
1361	Bt1Gc1708	Bt1G3371	3207-4163	g3892894	304	387	7.40E-36	29	100	(Y15477) hypothetical protein [Staphylococcus aureus]
1361	Bt1Gc1708	Bt1G3372	4678-6021	g2632444	1857	1797	2.90E-185	79	100	(Z99104) similar to phosphoglucomutase (glycolysis) [Bacillus subtilis]
1361	Bt1Gc1708	Bt1G3373	6511-8310	g1169919	2277	2209	6.30E-229	74	100	GLUCOSAMINE--FRUCTOSE-6-PHOSPHATE AMINOTRANSFERASE (ISOMERIZING) (HEXOSEPHOSPHATE AMINOTRANSFERASE) (D-FRUCTOSE-6-PHOSPHATE AMIDOTRANSFERASE) (GFAT) (L-GLUTAMINE-D-FRUCTOSE-6-PHOSPHATE AMIDOTRANSFERASE) (GLUCOSAMINE-6-PH...
1361	Bt1Gc1708	Bt1G3374	8601-10517	g2635720	420	411	6.30E-59	26	100	(Z99120) similar to acylaminoacyl-peptidase [Bacillus subtilis]
1361	Bt1Gc1708	Bt1G3375	12205-12978	g2633352	653	724	1.40E-71	54	100	(Z99110) alternate gene name: yidA; similar to 3-oxoacyl-acyl-carrier protein reductase [Bacillus subtilis]
1361	Bt1Gc1708	Bt1G3376	13670-13020	g401471	447	236	2.50E-24	42	100	HYPOTHETICAL ABC TRANSPORTER PERMEASE

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1361	Bt1Gc1708	Bt1G3377	14192-13690	g2314761	223	298	2.00E-26	38	49	PROTEIN YAE [Escherichia coli] (AE000655) ABC transporter, ATP-binding protein (abc) [Helicobacter pylori 26695]
1362	Bt1Gc1707	Bt1G3378	384-1	g2497382	616	622	9.30E-61	95	29	TRANSPORSE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1362	Bt1Gc1707	Bt1G3379	470-720	g1673402	229	277	1.10E-23	51	19	(Z82044) hypothetical 54.4 kd protein [Bacillus subtilis]
1362	Bt1Gc1707	Bt1G3380	2215-1679	g80034	337	350	6.20E-32	38	99	hypothetical protein 176 (SAGP 5' region) - Streptococcus pyogenes []
1362	Bt1Gc1707	Bt1G3381	3529-2588	g3687419	969	841	5.80E-84	60	100	(Y17554) carbanate kinase [Bacillus licheniformis]
1362	Bt1Gc1707	Bt1G3382	4982-3570	g3687418	1700	1450	1.70E-148	69	100	(Y17554) permease [Bacillus licheniformis]
1362	Bt1Gc1707	Bt1G3383	6077-5085	g3687417	1320	1298	2.20E-132	73	100	(Y17554) ornithine carbamoyltransferase [Bacillus licheniformis]
1362	Bt1Gc1707	Bt1G3384	7349-6111	g3687416	1563	1570	3.20E-161	69	100	(Y17554) arginine deiminase [Bacillus licheniformis]
1362	Bt1Gc1707	Bt1G3385	8061-7615	g4127534	336	295	4.20E-26	46	99	(AJ010954) arginine repressor [Bacillus stearothermophilus]
1362	Bt1Gc1707	Bt1G3386	11214-8211	g4512389	450	298	5.60E-34	27	58	(AB011838) methyl-accepting chemotaxis protein [Bacillus halodurans]
1362	Bt1Gc1707	Bt1G3387	9706-10432	g2649950	657	639	1.50E-62	56	100	(AE001058) glutamine ABC transporter, ATP-binding protein (glnQ) [Archaeoglobus fulgidus]
1362	Bt1Gc1707	Bt1G3388	10247-12148	g1652664	502	636	3.10E-62	33	100	(D90907) glutamine-binding periplasmic protein [Synechocystis sp.]
1362	Bt1Gc1707	Bt1G3389	13284-11512	g1652983	262	376	1.10E-34	27	100	(D90910) hypothetical protein [Synechocystis sp.]
1362	Bt1Gc1707	Bt1G3390	13869-14976	g2226259	600	591	1.80E-57	47	50	(Y14084) hypothetical protein [Bacillus subtilis]
1362	Bt1Gc1707	Bt1G3391	14636-14976	g225559	549	565	1.00E-54	98	24	ORF IS231C [Bacillus thuringiensis]
1363	Bt1Gc1711	Bt1G3392	1-1035	g2983197	792	843	3.50E-84	48	70	(AE000696) anthranilate synthase component I [Aquifex aeolicus]
1363	Bt1Gc1711	Bt1G3393	877-2606	g586124	1132	1133	6.60E-115	45	100	ANTHRANILATE SYNTHASE COMPONENT II (CONTAINS: GLUTAMINE AMIDOTRANSFERASE; ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE [Thermotoga maritima])
1363	Bt1Gc1711	Bt1G3394	2626-3380	g2499065	415	465	4.00E-44	42	96	INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE (IGPS) [Bacillus megaterium]
1363	Bt1Gc1711	Bt1G3395	3986-5191	g267168	1429	1438	3.20E-147	66	100	TRYPTOPHAN SYNTHASE BETA CHAIN [Lactococcus lactis]
1363	Bt1Gc1711	Bt1G3396	5189-5947	g267167	511	440	1.80E-41	40	100	TRYPTOPHAN SYNTHASE ALPHA CHAIN

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1363	Bt1Gc1711	Bt1G3397	8333-8771	g1665854	301	303	5.90E-27	46	99	[Lactococcus lactis] (Z82015) yukJ [Bacillus subtilis]
1363	Bt1Gc1711	Bt1G3398	12532-13866	g2688656	880	918	4.00E-92	39	100	(AE001172) NADH oxidase, water-forming (nox) [Borrelia burgdorferi]
1364	Bt1Gc1709	Bt1G3399	1489-212	g2226247	1156	1159	1.20E-117	51	100	(Y14083) hypothetical protein [Bacillus subtilis]
1364	Bt1Gc1709	Bt1G3400	3338-4115	g2635819	650	706	1.20E-69	55	71	(Z99120) similar to two-component response regulator [Bacillus subtilis]
1364	Bt1Gc1709	Bt1G3401	4131-5847	g2635818	1068	1066	8.30E-108	38	100	(Z99120) similar to two-component sensor histidine kinase [Bacillus subtilis]
1364	Bt1Gc1709	Bt1G3402	7419-6588	g3582223	331	460	1.40E-43	36	100	(AE001272) positive transcriptional activator [Lactococcus lactis]
1364	Bt1Gc1709	Bt1G3403	7641-8124	g2226208	244	316	2.50E-28	40	99	(Y14082) hypothetical protein [Bacillus subtilis]
1364	Bt1Gc1709	Bt1G3404	7678-8180	g3688548	86	185	1.90E-14	28	81	(A010320) RNA polymerase sigma factor [Streptomyces coelicolor]
1364	Bt1Gc1709	Bt1G3405	10265-9291	g1945094	999	1001	6.40E-101	55	100	(D88802) ydhJ [Bacillus subtilis]
1364	Bt1Gc1709	Bt1G3406	11508-10394	g2622286	197	319	1.20E-28	29	100	(AE000886) conserved protein [Methanobacterium thermoautotrophicum]
1364	Bt1Gc1709	Bt1G3407	12821-8950	g2496971	137	201	1.20E-14	27	55	HYPOTHETICAL 69.0 KD PROTEIN F40B5.2 IN CHROMOSOME X [Caenorhabditis elegans]
1364	Bt1Gc1709	Bt1G3408	12577-11656	g1742172	537	613	8.40E-60	42	100	(D90770) ORF_ID:o260#11; similar to [Escherichia coli]
1364	Bt1Gc1709	Bt1G3409	13556-15849	g1930114	486	541	3.60E-52	39	78	(U94743) hemolysin II [Bacillus cereus]
1365	Bt1Gc1712	Bt1G3410	1-1203	g730217	1617	1628	2.30E-167	75	100	ORNITHINE AMINOTRANSFERASE (ORNITHINE--OXO-ACID AMINOTRANSFERASE) [Bacillus subtilis]
1365	Bt1Gc1712	Bt1G3411	2529-1633	g2145376	869	804	4.80E-80	54	100	(Y09476) YisK [Bacillus subtilis]
1365	Bt1Gc1712	Bt1G3412	2776-2952	g2984717	289	289	1.80E-25	93	98	(AF053927) YisI homolog [Bacillus cereus]
1365	Bt1Gc1712	Bt1G3413	3052-3270	g2984718	368	368	7.70E-34	99	99	(AF053927) probably spore germination protein A [Bacillus cereus]
1365	Bt1Gc1712	Bt1G3414	3288-3491	g2984719	327	247	5.10E-21	99	99	(AF053927) probable spore germination protein B [Bacillus cereus]
1365	Bt1Gc1712	Bt1G3415	3562-4173	g2984720	1047	948	2.70E-95	99	100	(AF053927) probable spore germination protein C [Bacillus cereus]
1365	Bt1Gc1712	Bt1G3416	4183-4374	g2984721	320	320	9.40E-29	100	98	(AF053927) probable spore germination protein D [Bacillus cereus]
1365	Bt1Gc1712	Bt1G3417	4393-4776	g2984722	656	656	2.30E-64	100	99	(AF053927) probable spore germination protein E [Bacillus cereus]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1365	Bt1Gc1712	Bt1G3418	4822-5034	g2984723	366	292	8.70E-26	100	99	(AF053927) probable spore germination protein F [Bacillus cereus]
1365	Bt1Gc1712	Bt1G3419	9194-5484	g113345	3258	3322	0	54	100	ATP-DEPENDENT NUCLEASE SUBUNIT A [Bacillus subtilis]
1365	Bt1Gc1712	Bt1G3420	12712-9191	g113346	2954	2990	1.1e-311	50	100	ATP-DEPENDENT NUCLEASE SUBUNIT B [Bacillus subtilis]
1365	Bt1Gc1712	Bt1G3421	13305-12841	g729934	543	546	1.10E-52	63	85	SIGNAL PEPTIDASE I (SPASE I) (LEADER PEPTIDASE I) [Bacillus caldolyticus]
1366	Bt1Gc1713	Bt1G3422	1021-1470	g730784	459	468	1.90E-44	57	99	STAGE V SPORULATION PROTEIN AC [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3423	1474-2490	g730785	805	831	6.60E-83	48	100	STAGE V SPORULATION PROTEIN AD [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3424	2490-2843	g730786	335	385	1.20E-35	57	37	STAGE V SPORULATION PROTEIN AE [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3425	3091-3792	g1881360	348	370	4.70E-34	34	100	(AB001488) SIMILAR TO YDFR GENE PRODUCT OF THIS ENTRY (YDFR_BACSU). [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3426	4189-4779	g3287962	877	883	2.00E-88	88	99	ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT (ENDOPEPTIDASE CLP) (CASEINOLYTIC PROTEASE) (PROTEASE TI) [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3427	5346-5092	g1945653	293	259	2.70E-22	67	99	(Z94043) hypothetical protein [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3428	6320-5373	g1945652	1320	1320	1.00E-134	80	100	(Z94043) hypothetical protein [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3429	7406-6406	g2635989	1063	1060	3.60E-107	64	100	(Z99121) similar to hypothetical proteins [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3430	8252-7369	g2811049	1026	1049	5.30E-106	68	100	HYPOTHETICAL 33.9 KD PROTEIN IN CRH-TRXB INTERGENIC REGION [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3431	8734-8261	g1945649	578	592	1.40E-57	66	99	(Z94043) hypothetical protein [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3432	9755-8952	g2688692	182	148	4.00E-10	29	100	(AE001175) B. burgdorferi predicted coding region BB0759 [Borrelia burgdorferi]
1366	Bt1Gc1713	Bt1G3433	10883-9936	g3123300	1411	1411	2.30E-144	84	100	THIOREDOXIN REDUCTASE (GENERAL STRESS PROTEIN 35) (GSP35) [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3434	12480-10972	g1945645	692	745	8.60E-74	36	100	(Z94043) hypothetical protein [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3435	13132-12619	g2618861	709	725	1.10E-71	78	99	(AF017113) putative acetyltransferase [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3436	13825-13178	g2618860	717	662	5.40E-65	63	100	(AF017113) HP(Ser-P) phosphatase [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3437	14693-13890	g3122350	972	807	2.30E-80	67	100	PROLIPROTEIN DIACYLGLYCERYL TRANSFERASE (SPORE GERMINATION PROTEIN GERF) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvr	NCBI gi description
1366	Bt1Gc1713	Bt1G3438	15652-14723	g3122218	1211	1153	5.00E-117	75	100	HPR(SER) KINASE [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3439	16192-15821	g2618847	311	202	3.00E-16	51	99	(AF017113) YvID [Bacillus subtilis]
1366	Bt1Gc1713	Bt1G3440	18080-16814	g3915204	1815	1836	2.10E-189	82	44	EXCINUCLEASE ABC SUBUNIT A [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3441	1146-3297	g4490609	2358	2253	1.40E-233	64	100	(AJ133495) ribonucleotide reductase major subunit [Staphylococcus aureus]
1367	Bt1Gc1714	Bt1G3442	3443-4873	g1710395	892	952	1.00E-95	57	100	RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE BETA CHAIN (RIBONUCLEOTIDE REDUCTASE) [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3443	5078-5440	g1724000	266	272	1.10E-23	42	99	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN PRK-CSPB INTERGENIC REGION [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3444	5461-6147	g2314379	495	460	1.40E-43	42	100	(AE000627) ABC transporter, ATP-binding protein (YhcG) [Helicobacter pylori 26695]
1367	Bt1Gc1714	Bt1G3445	6947-7863	g1881259	684	393	3.90E-60	46	100	(AB001488) PROBABLE TRANSPORT ATP BINDING PROTEIN. [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3446	11088-9393	g1881235	357	427	4.30E-40	44	77	(AB001488) PROBABLE TRANSCRIPTIONAL REGULATOR, SIMILAR TO LEU RESPONSIVE REGULATOR. [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3447	10372-10935	g1934645	470	482	6.40E-46	52	99	(U93876) hypothetical protein YrdC [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3448	12125-11440	g2632018	442	494	3.40E-47	43	100	(AJ002571) YkfA [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3449	12481-13524	g1934639	588	587	4.80E-57	40	100	(U93875) 2-nitropropane dioxygenase [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3450	14760-13573	g729343	1013	1020	6.20E-103	48	100	PROTEIN DLT D PRECURSOR [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3451	14993-14760	g1405337	288	289	1.80E-25	69	99	(D86240) D-alanyl carrier protein [Staphylococcus aureus]
1367	Bt1Gc1714	Bt1G3452	16240-15065	g729341	1164	1080	2.70E-109	59	100	DLTB PROTEIN [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3453	17751-16240	g729340	1435	1460	1.50E-149	56	100	D-ALANINE-ACTIVATING ENZYME (DAE) (D-ALANINE-D-ALANYL CARRIER PROTEIN LIGASE) (DCL) [Bacillus subtilis]
1367	Bt1Gc1714	Bt1G3454	18245-19347	g584749	893	896	8.60E-90	51	100	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (AMINOACYLASE) [Bacillus stearothermophilus]
1367	Bt1Gc1714	Bt1G3455	19817-20665	g2632237	350	323	4.50E-29	44	99	(AJ222587) YkuP protein [Bacillus subtilis]
1368	Bt1Gc1706	Bt1G3456	1831-74	g585371	2029	1684	2.70E-173	70	100	PYRUVATE KINASE (PK) [Bacillus stearothermophilus]
1368	Bt1Gc1706	Bt1G3457	2841-1885	g125125	1422	1320	1.00E-134	86	100	6-PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE) [Bacillus stearothermophilus]
1368	Bt1Gc1706	Bt1G3458	3947-2973	g3121736	1245	1097	4.30E-111	72	100	ACETYL-COENZYME A CARBOXYLASE CARBOXYL

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1368	Bt1Gc1706	Bt1G3459	4717-3932	g2293262	971	979	1.40E-98	73	100	TRANSFERASE SUBUNIT ALPHA [Bacillus subtilis]
1368	Bt1Gc1706	Bt1G3460	5577-3921	g1945709	142	183	3.10E-14	38	44	(AF008220) acetyl-CoA carboxylase subunit [Bacillus subtilis]
1368	Bt1Gc1706	Bt1G3461	6881-5652	g2293261	1727	1589	3.10E-163	83	100	(Z94043) hypothetical protein [Bacillus subtilis]
1368	Bt1Gc1706	Bt1G3462	10314-6985	g2293260	2861	2951	1.50E-307	51	100	(AF008220) YtsJ [Bacillus subtilis]
1368	Bt1Gc1706	Bt1G3463	10767-11267	g2293326	296	243	1.40E-20	36	99	(AF008220) DNA-polymerase III alpha-chain [Bacillus subtilis]
1368	Bt1Gc1706	Bt1G3464	12254-11316	g2293259	1029	1035	1.60E-104	61	100	(AF008220) YtrI [Bacillus subtilis]
1368	Bt1Gc1706	Bt1G3465	12414-12713	g2293325	203	205	1.40E-16	42	99	(AF008220) YtqI [Bacillus subtilis]
1368	Bt1Gc1706	Bt1G3466	14514-13201	g2293258	1567	1586	6.50E-163	68	100	(AF008220) YtpI [Bacillus subtilis]
1368	Bt1Gc1706	Bt1G3467	15643-14325	g2649315	468	538	7.40E-52	46	100	(AF008220) Ytol [Bacillus subtilis]
1368	Bt1Gc1706	Bt1G3468	15796-16884	g2633757	1135	1141	9.40E-116	58	100	(AE001017) conserved hypothetical protein [Archaeoglobus fulgidus]
1369	Bt1Gc1715	Bt1G3469	392-63	g1172708	297	304	4.60E-27	55	99	(Z99111) similar to Xaa-Pro dipeptidase [Bacillus subtilis]
1369	Bt1Gc1715	Bt1G3470	1881-174	g1172710	933	978	1.80E-98	44	100	PTS SYSTEM, CELLOBIOSE-SPECIFIC IIA COMPONENT (EIIA-CEL) (CELLOBIOSE-PERMEASE IIA COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, A COMPONENT) (EIII-CEL) [Bacillus subtilis]
1369	Bt1Gc1715	Bt1G3471	2200-1901	g2499992	316	316	2.50E-28	60	99	PTS SYSTEM, CELLOBIOSE-SPECIFIC IIC COMPONENT (EIIC-CEL) (CELLOBIOSE-PERMEASE IIC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, C COMPONENT) [Bacillus subtilis]
1369	Bt1Gc1715	Bt1G3472	4184-2838	g2499991	1515	1394	1.50E-142	65	100	PTS SYSTEM, CELLOBIOSE-SPECIFIC IIB COMPONENT (EIIB-CEL) (CELLOBIOSE-PERMEASE IIB COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, B COMPONENT) [Bacillus subtilis]
1369	Bt1Gc1715	Bt1G3473	4492-4187	g1172709	332	343	3.40E-31	62	99	PTS SYSTEM, CELLOBIOSE-SPECIFIC IIB COMPONENT (EIIB-CEL) (CELLOBIOSE-PERMEASE IIB COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, C COMPONENT) [Bacillus subtilis]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1369	Bt1Gc1715	Bt1G3474	4711-5952	g121467	328	370	4.70E-34	25	100	IIB COMPONENT (PHOSPHOTRANSFERASE ENZYME II, B COMPONENT) [Bacillus subtilis]
1369	Bt1Gc1715	Bt1G3475	7279-7650	g2633029	150	210	4.20E-17	35	99	PROTON/SODIUM-GLUTAMATE SYMPORT PROTEIN (GLUTAMATE-ASPARTATE CARRIER PROTEIN) [Bacillus stearothermophilus]
1369	Bt1Gc1715	Bt1G3476	9691-7700	g2851530	766	601	1.60E-58	32	100	(Z99107) similar to hypothetical proteins [Bacillus subtilis]
1369	Bt1Gc1715	Bt1G3477	9835-10716	g4757086	502	523	2.90E-50	38	100	HYPOTHETICAL 84.1 KD PROTEIN IN HEMY-GLTT INTERGENIC REGION (ORFB) [Bacillus subtilis]
1369	Bt1Gc1715	Bt1G3478	12816-10764	g2632727	455	511	5.40E-49	35	100	(AL049763) putative transcriptional regulator [Streptomyces coelicolor]
1369	Bt1Gc1715	Bt1G3479	12363-13690	g2293293	1173	1184	2.60E-120	54	100	(Z99106) ydaJ [Bacillus subtilis]
1369	Bt1Gc1715	Bt1G3480	13688-14722	g2293292	814	696	1.30E-68	46	100	(AF008220) YthA [Bacillus subtilis]
1369	Bt1Gc1715	Bt1G3481	15178-16506	g1929338	1337	1136	3.20E-115	58	100	(AF008220) YthB [Bacillus subtilis]
1369	Bt1Gc1715	Bt1G3482	18377-16617	g3041750	2468	2439	2.70E-253	78	100	(Z93767) similar to ArsB integral membrane protein presumably containing the arsenate-arsenite-antimony transmembrane channel [Bacillus subtilis]
1370	Bt1Gc1717	Bt1G3483	1332-382	g80325	785	746	6.70E-74	46	100	THIAMIN BIOSYNTHESIS PROTEIN THIC [Bacillus subtilis]
1370	Bt1Gc1717	Bt1G3484	3989-3291	g2619046	739	681	5.20E-67	59	100	L-lactate dehydrogenase (EC 1.1.1.27) - Bacillus subtilis []
1370	Bt1Gc1717	Bt1G3485	5913-4132	g2635778	2319	2093	1.20E-216	76	100	(AF027868) YobT [Bacillus subtilis]
1370	Bt1Gc1717	Bt1G3486	7430-6261	g2635779	1564	1574	1.20E-161	79	100	(Z99120) similar to butyryl-CoA dehydrogenase [Bacillus subtilis]
1370	Bt1Gc1717	Bt1G3487	9910-7455	g2635780	2861	2800	1.50E-291	68	100	(Z99120) similar to acetyl-CoA C-acyltransferase [Bacillus subtilis]
1370	Bt1Gc1717	Bt1G3488	10321-12466	g1303997	267	302	7.60E-27	34	74	(Z99120) similar to 3-hydroxyacyl-CoA dehydrogenase [Bacillus subtilis]
1370	Bt1Gc1717	Bt1G3489	12073-11168	g2635781	974	974	4.70E-98	60	100	(D84432) DacF [Bacillus subtilis]
1370	Bt1Gc1717	Bt1G3490	12681-12397	g2635789	283	254	9.20E-22	58	99	(Z99120) similar to proline dehydrogenase [Bacillus subtilis]
1370	Bt1Gc1717	Bt1G3491	12990-12676	g3023656	271	292	8.70E-26	54	99	(Z99120) yusU [Bacillus subtilis]
1370	Bt1Gc1717	Bt1G3492	15068-13077	g1708962	756	753	1.20E-74	31	100	HYPOTHETICAL 12.1 KD PROTEIN IN GLKA 3'REGION (ORF2) [Staphylococcus xylosus]
1370	Bt1Gc1717	Bt1G3493	15165-13077	g2626835	427	524	2.30E-50	27	83	METHYL-ACCEPTING CHEMOTAXIS PROTEIN MCPC [Bacillus subtilis]
1370	Bt1Gc1717	Bt1G3493	15165-13077	g2626835	427	524	2.30E-50	27	83	(D86947) chemotactic transducer [Pseudomonas]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1371	Bt1Gc1720	Bt1G3494	842-1	g2239276	1101	834	2.00E-121	72	68	aeruginosa] neutral protease A [Bacillus thuringiensis] (L77763)
1371	Bt1Gc1720	Bt1G3495	2546-1275	g473955	2128	2046	1.20E-211	97	100	DNA-binding protein [Lactobacillus sp.] (D29674)
1371	Bt1Gc1720	Bt1G3496	2739-4208	g2635670	1842	1869	6.70E-193	70	100	similar to nicotinate phosphoribosyltransferase [Bacillus subtilis] (Z99120)
1371	Bt1Gc1720	Bt1G3497	6772-4415	g2635862	2322	2306	3.30E-239	66	100	similar to heavy metal-transporting ATPase [Bacillus subtilis] (Z99121)
1371	Bt1Gc1720	Bt1G3498	7199-6804	g2495428	247	262	1.30E-22	36	99	TRANSCRIPTIONAL REPRESSOR SMTB HOMOLOG [Synechocystis sp.] (D86418)
1371	Bt1Gc1720	Bt1G3499	9489-8110	g2116759	1328	1006	1.90E-101	52	100	YfnA [Bacillus subtilis] (D86418)
1371	Bt1Gc1720	Bt1G3500	10723-9611	g118515	1400	1409	3.70E-144	73	100	ALANINE DEHYDROGENASE [Bacillus sphaericus] (Y14082)
1371	Bt1Gc1720	Bt1G3501	11884-11189	g2226218	201	261	1.70E-22	28	100	hypothetical protein [Bacillus subtilis]
1371	Bt1Gc1720	Bt1G3502	13602-13123	g3915537	354	320	9.40E-29	45	99	HYPOTHETICAL 17.5 KD PROTEIN IN SIGV-GREA INTERGENIC REGION [Bacillus subtilis] (Z99117)
1371	Bt1Gc1720	Bt1G3503	16554-13618	g2635168	3878	3768	0	72	100	similar to formate dehydrogenase [Bacillus subtilis]
1371	Bt1Gc1720	Bt1G3504	16849-17640	g730100	693	735	9.90E-73	53	100	NARQ PROTEIN [Bacillus subtilis]
1372	Bt1Gc1716	Bt1G3505	5032-1783	g2633696	863	954	6.10E-96	37	74	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
1372	Bt1Gc1716	Bt1G3506	10436-12289	g400640	1061	927	8.50E-112	40	100	NADH-PLASTOQUINONE OXIDOREDUCTASE CHAIN 5 [Synechococcus sp.] (AJ235273)
1372	Bt1Gc1716	Bt1G3507	11328-15150	g3861319	399	425	1.60E-39	37	53	NADH DEHYDROGENASE I CHAIN L (nuoL1) [Rickettsia prowazekii] (Y14080)
1373	Bt1Gc1718	Bt1G3508	1530-307	g2226151	1577	1166	2.10E-118	77	100	hypothetical protein [Bacillus subtilis]
1373	Bt1Gc1718	Bt1G3509	2022-1531	g2226150	511	523	2.90E-50	62	99	hypothetical protein [Bacillus subtilis]
1373	Bt1Gc1718	Bt1G3510	5601-4843	g2293177	643	673	3.70E-66	49	100	(AF008220) transporter [Bacillus subtilis]
1373	Bt1Gc1718	Bt1G3511	6985-5951	g418336	993	841	5.80E-84	57	100	HYPOTHETICAL 39.9 KD PROTEIN IN AMYLASE 3'REGION [Bacillus acidopullulyticus] (AF008220)
1373	Bt1Gc1718	Bt1G3512	7662-8627	g2293302	1344	1354	2.50E-138	77	100	YtqA [Bacillus subtilis]
1373	Bt1Gc1718	Bt1G3513	8618-9202	g2293301	564	583	1.30E-56	57	99	YtqB [Bacillus subtilis]
1373	Bt1Gc1718	Bt1G3514	9547-10209	g1945684	624	653	4.80E-64	53	100	amino acid racemase [Bacillus subtilis]
1373	Bt1Gc1718	Bt1G3515	10365-11693	g1498192	300	256	4.80E-33	26	100	putative [Pseudomonas aeruginosa] (D28119)
1373	Bt1Gc1718	Bt1G3516	10365-11693	g1498192	300	256	4.80E-33	26	100	putative [Pseudomonas aeruginosa] (D28119)
1373	Bt1Gc1718	Bt1G3517	13794-12695	g2293168	1153	1101	1.60E-111	58	100	YtpB [Bacillus subtilis] (AF008220)
1373	Bt1Gc1718	Bt1G3518	14638-13862	g2293167	735	735	9.90E-73	52	100	probable lysophospholipase [Bacillus subtilis] (AF008220)

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1373	Bt1Gc1718	Bt1G3519	14713-15225	g22933300	648	576	7.00E-56	75	99	(AF008220) YtoA [Bacillus subtilis]
1373	Bt1Gc1718	Bt1G3520	15222-17057	g1001478	568	657	1.80E-64	35	100	(D63999) hypothetical protein [Synechocystis sp.]
1373	Bt1Gc1718	Bt1G3521	17938-17335	g2633797	305	337	1.50E-30	35	99	(Z99111) molybdopterin-guanine dinucleotide biosynthesis protein A [Bacillus subtilis]
1373	Bt1Gc1718	Bt1G3522	18447-17938	g2293241	551	555	1.20E-53	61	99	(AF008220) putative molibdenum cofactor biosynthesis protein [Bacillus subtilis]
1373	Bt1Gc1718	Bt1G3523	20714-19636	g2293164	1525	1501	6.70E-154	84	90	(AF008220) SAM synthase [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3524	443-779	g732365	322	340	7.10E-31	64	99	HYPOTHETICAL 12.0 KD PROTEIN IN UNG-ROCA INTERGENIC REGION [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3525	1395-847	g732366	259	242	1.70E-20	40	99	HYPOTHETICAL 20.3 KD PROTEIN IN UNG-ROCA INTERGENIC REGION [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3526	1722-1294	g1175629	470	380	4.10E-35	61	99	CELL WALL HYDROLASE CWLJ [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3527	3187-1881	g461913	879	974	4.70E-98	46	100	D-ALANYL-D-ALANINE CARBOXYPEPTIDASE PRECURSOR (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) (CPASE) (PBP5) [Bacillus stearothermophilus]
1374	Bt1Gc1719	Bt1G3528	4128-3370	g732376	932	941	1.50E-94	67	100	HYPOTHETICAL 29.5 KD PROTEIN IN ROCC-PTA INTERGENIC REGION [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3529	4366-5334	g730415	1278	1278	2.80E-130	79	100	PROBABLE PHOSPHATE ACETYLTRANSFERASE (PHOSPHOTRANSACETYLASE) (VEGETATIVE PROTEIN 43) (VEG43) [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3530	6234-5398	g732378	845	861	4.40E-86	61	100	HYPOTHETICAL 31.4 KD PROTEIN IN PTA 3'REGION [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3531	7055-6363	g732380	574	402	1.90E-60	54	100	HYPOTHETICAL 29.7 KD PROTEIN IN PTA 3'REGION [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3532	7337-7558	g2636297	321	321	7.30E-29	77	99	(Z99123) ywzC [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3533	6884-8922	g1706795	963	957	3.00E-96	51	100	FERRICHROME TRANSPORT PERMEASE PROTEIN FHUB [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3534	8931-9938	g1706797	954	693	2.80E-68	54	100	FERRICHROME TRANSPORT PERMEASE PROTEIN FHUG [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3535	9957-10763	g1706796	922	871	3.80E-87	65	100	FERRICHROME TRANSPORT ATP-BINDING PROTEIN FHUC [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3536	10962-11897	g2633168	205	332	5.00E-30	28	100	(Z99108) similar to iron(III) dicitrate transport permease [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3537	12043-13341	g2506579	1659	1666	2.20E-171	71	100	HYPOTHETICAL 51.0 KD PROTEIN IN PTA 3'REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1374	Bt1Gc1719	Bt1G3538	13572-13387	g1565237	238	243	1.40E-20	79	98	[Bacillus subtilis] (Z80360) Unknown, highly similar to Pseudomonas putida 4-oxalocrotonate tautomerase [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3539	13685-14347	g1565238	620	552	2.40E-53	53	100	(Z80360) Unknown [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3540	14382-14897	g1565239	690	654	3.80E-64	76	99	(Z80360) Unknown [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3541	16978-15047	g1565240	1740	1741	2.50E-179	53	100	(Z80360) unknown, highly similar to penicillin binding protein [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3542	17799-17089	g1731058	434	453	7.50E-43	38	100	HYPOTHETICAL 27.0 KD PROTEIN IN SPO0A-MMGA INTERGENIC REGION [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3543	17952-18356	g2116754	260	283	7.80E-25	39	99	(D86418) YfmP [Bacillus subtilis]
1374	Bt1Gc1719	Bt1G3544	18365-19456	g4914624	1071	920	2.50E-92	57	92	(AJ009627) multidrug resistance transporter [Listeria monocytogenes]
1375	Bt1Gc1721	Bt1G3545	3447-4426	g4530143	383	354	2.30E-32	41	100	(AF085222) putative scaffolding protein [Streptococcus thermophilus bacteriophage DT1]
1375	Bt1Gc1721	Bt1G3546	6889-7482	g4126620	187	207	8.80E-17	31	100	(AB016282) ORF34 [bacteriophage phi-105]
1375	Bt1Gc1721	Bt1G3547	8579-11745	g3947462	234	346	1.30E-30	25	57	(AJ006589) gp43 [Bacteriophage phi-C31]
1375	Bt1Gc1721	Bt1G3548	11990-13839	g4126623	200	230	3.20E-19	32	79	(AB016282) ORF37 [bacteriophage phi-105]
1376	Bt1Gc1722	Bt1G3549	156-1067	g1731090	830	715	1.30E-70	53	100	HYPOTHETICAL 34.6 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3550	2475-3557	g132565	1005	1006	1.90E-101	54	100	RIBOFLAVIN-SPECIFIC DEAMINASE [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3551	3566-4204	g132633	656	618	2.50E-60	61	100	RIBOFLAVIN SYNTHASE ALPHA CHAIN [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3552	4223-5413	g121062	1532	1541	3.80E-158	73	100	GTP CYCLOHYDROLASE II / 3,4-DIHYDROXY-2-BUTANONE 4-PHOSPHATE SYNTHASE (DHP SYNTHASE) [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3553	5432-5893	g2497762	588	416	6.30E-39	73	99	6,7-DIMETHYL-8-RIBITYLLUMAZINE SYNTHASE (DMRL SYNTHASE) (LUMAZINE SYNTHASE) (RIBOFLAVIN SYNTHASE BETA CHAIN) [Bacillus amyloqueluctans]
1376	Bt1Gc1722	Bt1G3554	7401-6399	g1705464	1012	1031	4.20E-104	59	100	BIOTIN SYNTHASE (BIOTIN SYNTHETASE) [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3555	10108-8939	g115009	776	789	1.90E-78	43	100	8-AMINO-7-OXONONANOATE SYNTHASE (7-KETO-8-AMINO-PELARGONIC ACID SYNTHETASE) (7-KAP SYNTHETASE) (L-ALANINE--PIMELYL COA LIGASE) [Bacillus sphaericus]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1376	Bt1Gc1722	Bt1G3556	10814-10119	g1705466	495	520	6.00E-50	44	100	DETHIOBIOTIN SYNTHETASE (DETHIOBIOTIN SYNTHASE) (DTB SYNTHETASE) (DTBS) [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3557	12163-10814	g1705460	1416	1351	5.20E-138	59	100	ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE) [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3558	13315-14694	g3287970	1001	918	4.00E-92	41	100	NA(+)/H(+) ANTIPORTER [Bacillus firmus]
1376	Bt1Gc1722	Bt1G3559	17072-15133	g2626826	1450	1663	1.30E-173	48	44	(D83967) YfkN [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3560	20682-19729	g129265	924	887	7.70E-89	58	100	ORNITHINE CARBAMOYLTRANSFERASE, ANABOLIC (OTCASE) [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3561	21852-20699	g543846	1032	969	1.60E-97	54	100	ACETYLORNITHINE AMINOTRANSFERASE (ACOAT) [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3562	22619-21852	g584767	543	521	4.70E-50	48	100	ACETYLGLUTAMATE KINASE (NAG KINASE) (AGK) (N-ACETYL-L-GLUTAMATE 5-PHOSPHOTRANSFERASE) [Bacillus stearothermophilus]
1376	Bt1Gc1722	Bt1G3563	23854-22634	g584771	1233	1251	2.10E-127	60	100	GLUTAMATE N-ACETYLTRANSFERASE (ORNITHINE ACETYLTRANSFERASE) (ORNITHINE TRANSACETYLASE) (OATASE) / AMINO-ACID ACETYLTRANSFERASE (N-ACETYLGLUTAMATE SYNTHASE) (AGS) [Bacillus stearothermophilus]
1376	Bt1Gc1722	Bt1G3564	24893-23859	g80340	967	870	4.90E-87	53	100	N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) - Bacillus subtilis [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3565	26125-25337	g1731078	730	694	2.20E-68	56	100	HYPOTHETICAL OXIDOREDUCTASE IN ANSR-BMRU INTERGENIC REGION [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3566	27097-25719	g1731077	706	761	1.70E-75	47	100	HYPOTHETICAL 35.8 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3567	27263-28099	g1709791	644	604	7.50E-59	48	100	PYRROLINE-5-CARBOXYLATE REDUCTASE HOMOLOG 2 [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3568	29022-28112	g2337815	1072	902	2.00E-90	72	100	(Y13937) putative YhaP protein [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3569	29700-29041	g2337814	671	649	1.30E-63	59	100	(Y13937) putative YhaQ protein [Bacillus subtilis]
1376	Bt1Gc1722	Bt1G3570	31495-30585	g1731073	964	989	1.20E-99	60	99	HYPOTHETICAL 34.0 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]
1377	Bt1Gc1723	Bt1G3571	647-1	g2622920	669	717	8.00E-71	64	63	(AE000933) dTDP-glucose 4,6-dehydratase

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1377	Bt1Ge1723	Bt1G3572	1205-648	g3608395	517	543	2.20E-52	56	99	[Methanobacterium thermoautotrophicum] (AF071085) dTDP-4-dehydrohamnose 3,5-epimerase [Enterococcus faecalis]
1377	Bt1Ge1723	Bt1G3573	1951-1220	g730818	668	644	4.30E-63	52	100	SPORE COAT POLYSACCHARIDE BIOSYNTHESIS PROTEIN SPSI [Bacillus subtilis]
1377	Bt1Ge1723	Bt1G3574	4015-1843	g3608402	159	232	4.60E-19	30	50	(AF071085) Orfde14 [Enterococcus faecalis]
1377	Bt1Ge1723	Bt1G3575	6500-7345	g303643	574	629	1.70E-61	46	100	(D16097) mycinamicinIII O-methyltransferase [Micromonospora griseorubida]
1377	Bt1Ge1723	Bt1G3576	7667-8281	g2829568	528	534	2.00E-51	47	100	HYPOTHETICAL 23.1 KD PROTEIN CY277.20C [Mycobacterium tuberculosis]
1377	Bt1Ge1723	Bt1G3577	8298-8649	g1651989	166	230	3.60E-18	40	17	(D90901) hypothetical protein [Synecocystis sp.]
1377	Bt1Ge1723	Bt1G3578	9451-10617	g4887204	337	482	6.40E-46	32	100	(AF147448) rod-shape-determining protein [Pseudomonas aeruginosa]
1377	Bt1Ge1723	Bt1G3579	10697-11431	g2633517	753	771	1.50E-76	57	100	(Z99110) similar to diadenosine tetraphosphatase [Bacillus subtilis]
1377	Bt1Ge1723	Bt1G3580	12074-13384	g4126627	100	299	2.10E-25	30	31	(AB016282) ORF41 [bacteriophage phi-105]
1377	Bt1Ge1723	Bt1G3581	13843-12995	g3915487	803	813	5.40E-81	54	100	HYPOTHETICAL 31.5 KD PROTEIN IN MECA-TENA INTERGENIC REGION [Bacillus subtilis]
1377	Bt1Ge1723	Bt1G3582	14703-13906	g2633515	1046	946	4.30E-95	72	100	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
1377	Bt1Ge1723	Bt1G3583	15363-14728	g2633514	707	721	3.00E-71	65	100	(Z99110) similar to GTP pyrophosphokinase [Bacillus subtilis]
1377	Bt1Ge1723	Bt1G3584	15762-15394	g2633513	210	157	1.80E-11	39	99	(Z99110) yjbl [Bacillus subtilis]
1377	Bt1Ge1723	Bt1G3585	15907-16479	g2633512	429	428	3.40E-40	50	99	(Z99110) yjbk [Bacillus subtilis]
1377	Bt1Ge1723	Bt1G3586	16653-17051	g2633510	438	458	2.20E-43	67	99	(Z99110) yjbl [Bacillus subtilis]
1377	Bt1Ge1723	Bt1G3587	17126-17956	g2633509	591	523	2.90E-50	43	100	(Z99110) yjbH [Bacillus subtilis]
1377	Bt1Ge1723	Bt1G3588	23082-18475	g1651216	2312	2182	4.60E-226	70	100	(D88209) Pz-peptidase [Bacillus licheniformis]
1377	Bt1Ge1723	Bt1G3589	23784-23307	g585473	299	333	3.90E-30	51	70	NEGATIVE REGULATOR OF GENETIC COMPETENCE MECA [Bacillus subtilis]
1378	Bt1Ge1724	Bt1G3590	1-369	g1749770	280	308	1.70E-27	52	41	(Y09946) transposase [Bacillus thuringiensis]
1378	Bt1Ge1724	Bt1G3591	3443-3196	g2128207	199	228	5.30E-19	57	99	hypothetical protein MJ0272 - Methanococcus jannaschii [Methanococcus jannaschii]
1378	Bt1Ge1724	Bt1G3592	4215-7040	g1071810	2494	2584	1.10E-268	53	100	oxoglutarate dehydrogenase (lipoamide) (EC 1.2.4.2) - Bacillus subtilis []
1378	Bt1Ge1724	Bt1G3593	7219-8475	g129041	1347	1235	1.00E-125	66	100	DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE COMPONENT OF 2-OXOGLUTARATE

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1378	Bt1Gc1724	Bt1G3594	11311-12462	g4218544	171	377	6.10E-34	30	50	DEHYDROGENASE COMPLEX (E2) [Bacillus subtilis]
1378	Bt1Gc1724	Bt1G3595	15151-14771	g2497400	187	229	4.10E-19	38	57	(AJ011500) gra-orf12 [Streptomyces violaceoruber]
1379	Bt1Gc1726	Bt1G3596	1-1338	g1934805	242	311	8.40E-28	39	43	HYPOTHETICAL TRANSPOSASE-LIKE PROTEIN HI1721 [Haemophilus influenzae Rd]
1379	Bt1Gc1726	Bt1G3597	1088-684	g1934804	316	339	9.10E-31	45	99	(Z93936) unknown [Bacillus subtilis]
1379	Bt1Gc1726	Bt1G3598	2545-1199	g120717	2051	2006	2.00E-207	85	100	(Z93936) unknown [Bacillus subtilis]
1379	Bt1Gc1726	Bt1G3599	2823-3053	g2635634	284	223	1.80E-18	74	99	GLUCOSE-6-PHOSPHATE ISOMERASE A (GPI A) (PHOSPHOGLUCOSE ISOMERASE A) [Bacillus stearothermophilus]
1379	Bt1Gc1726	Bt1G3600	3510-3161	g3123231	274	270	1.90E-23	52	99	(Z99120) yuzA [Bacillus subtilis]
1379	Bt1Gc1726	Bt1G3601	4951-3885	g1934791	1073	1088	3.90E-110	60	100	GENERAL STRESS PROTEIN 13 (GSP13) [Bacillus subtilis]
1379	Bt1Gc1726	Bt1G3602	5451-4954	g2125791	551	471	9.30E-45	62	99	(Z93934) unknown [Bacillus subtilis]
1379	Bt1Gc1726	Bt1G3603	5708-6562	g1945659	901	866	1.30E-86	61	100	(Z94043) hypothetical protein [Bacillus subtilis]
1379	Bt1Gc1726	Bt1G3604	6675-7502	g1934789	760	778	2.70E-77	53	100	(Z93934) unknown [Bacillus subtilis]
1379	Bt1Gc1726	Bt1G3605	8001-9146	g585644	913	929	2.70E-93	45	100	PUTATIVE AMINOTRANSFERASE B [Bacillus subtilis]
1379	Bt1Gc1726	Bt1G3606	9885-9379	g2982937	317	345	2.10E-31	37	99	(AE000679) superoxide dismutase (Cu/Zn) [Aquifex aeolicus]
1379	Bt1Gc1726	Bt1G3607	10301-10684	g585341	379	391	2.80E-36	59	99	KINASE-ASSOCIATED PROTEIN B [Bacillus subtilis]
1379	Bt1Gc1726	Bt1G3608	11345-10731	g1934782	636	636	3.10E-62	57	100	(Z93933) unknown [Bacillus subtilis]
1379	Bt1Gc1726	Bt1G3609	11657-12964	g3258354	973	878	6.90E-88	50	100	(AP000007) 424aa long hypothetical protein [Pyrococcus horikoshii]
1379	Bt1Gc1726	Bt1G3610	13497-11889	g4584088	711	754	9.60E-75	90	81	(AJ010128) hypothetical protein [Bacillus cereus]
1379	Bt1Gc1726	Bt1G3611	13615-14325	g4584089	1184	1184	2.60E-120	93	100	(AJ010128) DNA alkylation repair enzyme [Bacillus cereus]
1379	Bt1Gc1726	Bt1G3612	14675-15325	g4584090	1157	1157	1.90E-117	100	100	(AJ010128) glycyl-tRNA synthetase [Bacillus cereus]
1379	Bt1Gc1726	Bt1G3613	14669-16048	g2500998	401	507	1.50E-90	32	100	PROBABLE GLYCYL-TRNA SYNTHETASE (GLYCINE-TRNA LIGASE) (GLYRS) [Methanococcus jannaschii]
1380	Bt1Gc1728	Bt1G3614	4134-454	g1041115	892	993	4.50E-100	39	100	(D78016) TRAC [Enterococcus faecalis]
1380	Bt1Gc1728	Bt1G3615	5357-4530	g2635853	1093	1095	7.00E-111	74	100	(Z99121) alternate gene name: yvsB; similar to plant-metabolite dehydrogenase [Bacillus subtilis]
1380	Bt1Gc1728	Bt1G3616	6156-5331	g1709791	398	428	3.40E-40	33	100	PYRROLINE-5-CARBOXYLATE REDUCTASE HOMOLOG 2 [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1380	Bt1Gc1728	Bt1G3617	7302-6448	g1805461	935	820	9.70E-82	62	100	(D50453) homologues to hypothetical 30.5 kDa protein gdh1 5' region of <i>B. megaterium</i> [Bacillus subtilis]
1380	Bt1Gc1728	Bt1G3618	8175-7613	g2635852	538	506	1.80E-48	59	99	(Z99121) alternate gene name: yvsC; similar to molybdenum transport permease [Bacillus subtilis]
1380	Bt1Gc1728	Bt1G3619	9070-8289	g2635851	700	737	6.10E-73	57	100	(Z99121) alternate gene name: yvsD; similar to molybdate-binding protein [Bacillus subtilis]
1380	Bt1Gc1728	Bt1G3620	11458-10595	g1772644	1133	1056	9.50E-107	74	100	(U62055) orfR gene product [Bacillus subtilis]
1380	Bt1Gc1728	Bt1G3621	11586-12554	g1934659	1177	1088	3.90E-110	68	100	(U93876) hypothetical protein YrdR [Bacillus subtilis]
1380	Bt1Gc1728	Bt1G3622	13210-13806	g2226210	515	518	9.80E-50	52	99	(Y14082) hypothetical protein [Bacillus subtilis]
1380	Bt1Gc1728	Bt1G3623	14282-15256	g2633758	701	743	1.40E-73	44	100	(Z99111) similar to transcriptional regulator (LacI family) [Bacillus subtilis]
1380	Bt1Gc1728	Bt1G3624	15871-16707	g1783243	766	722	2.40E-71	50	100	(D83026) homologous to jojC gene product ( <i>B. subtilis</i> ; prf2111327a); hypothetical [Bacillus subtilis]
1380	Bt1Gc1728	Bt1G3625	17837-18685	g1789981	796	806	3.00E-80	52	100	(AE000433) IS150 putative transposase [Escherichia coli]
1381	Bt1Gc1727	Bt1G3626	1-1035	g1881240	569	596	5.30E-58	37	83	(AB001488) FUNCTION UNKNOWN, SIMILAR PRODUCT IN MANY BACTERIA. [Bacillus subtilis]
1381	Bt1Gc1727	Bt1G3627	6530-1675	g4835822	1317	1216	1.10E-123	54	100	(AF102174) glycine betaine transporter BetL [Listeria monocytogenes]
1381	Bt1Gc1727	Bt1G3628	10487-13255	g2636153	2310	2406	8.40E-250	50	100	(Z99122) similar to SNF2 helicase [Bacillus subtilis]
1381	Bt1Gc1727	Bt1G3629	18484-17288	g2633808	1012	776	4.50E-77	52	100	(Z99111) similar to hypothetical proteins [Bacillus subtilis]
1381	Bt1Gc1727	Bt1G3630	19161-18475	g2832800	718	729	4.30E-72	62	100	(AJ223978) putative ABC transporter, YvrO [Bacillus subtilis]
1381	Bt1Gc1727	Bt1G3631	20321-19164	g2633806	471	343	3.40E-31	29	100	(Z99111) similar to hypothetical proteins from <i>B. subtilis</i> [Bacillus subtilis]
1381	Bt1Gc1727	Bt1G3632	21925-20600	g585209	1011	929	2.70E-93	46	100	GLYCEROL-3-PHOSPHATE TRANSPORTER (G-3-P TRANSPORTER) (G-3-P PERMEASE) [Bacillus subtilis]
1382	Bt1Gc1725	Bt1G3633	11136-123	g2506097	809	827	1.80E-82	49	100	ATP-BINDING PROTEIN ABC [Escherichia coli]
1382	Bt1Gc1725	Bt1G3634	3973-2429	g730598	1990	1990	1.00E-205	73	100	1-PYRROLINE-5-CARBOXYLATE DEHYDROGENASE (P5C DEHYDROGENASE) [Bacillus subtilis]
1382	Bt1Gc1725	Bt1G3635	8217-6211	g2632976	2359	2287	3.40E-237	69	100	(Z99107) similar to DNA ligase [Bacillus subtilis]
1382	Bt1Gc1725	Bt1G3636	10485-8236	g2577965	2592	2590	2.70E-269	68	100	(Y15254) PcrA protein [Bacillus subtilis]
1382	Bt1Gc1725	Bt1G3637	11184-10498	g2577964	793	811	8.70E-81	63	100	(Y15254) PcrB protein [Bacillus subtilis]
1382	Bt1Gc1725	Bt1G3638	14452-13190	g131611	1376	1399	4.30E-143	63	100	PHOSPHORIBOSYLAMINE--GLYCINE LIGASE (GARS) (GLYCINAMIDE RIBONUCLEOTIDE SYNTHETASE) (PHOSPHORIBOSYLGLYCINAMIDE SYNTHETASE)



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1382	Bt1Gc1725	Bt1G3639	16413-14878	g131638	1999	2007	1.60E-207	76	100	[Bacillus subtilis] PHOSPHORIBOSYLAMINOIMIDAZOLECARBOXAMID E FORMYLTRANSFERASE (AICAR TRANSFORMYLASE) / IMP CYCLOHYDROLASE (INOSINICASE) (IMP SYNTHETASE) (ATIC) [Bacillus subtilis]
1382	Bt1Gc1725	Bt1G3640	18059-17022	g131624	1254	1169	1.00E-118	67	100	PHOSPHORIBOSYLFORMYLGLYCINAMIDINE CYCLO-LIGASE (AIRS) (PHOSPHORIBOSYL- AMINOIMIDAZOLE SYNTHETASE) (AIR SYNTHASE)
1382	Bt1Gc1725	Bt1G3641	21784-19559	g131646	2967	2949	2.40E-307	76	100	[Bacillus subtilis] PHOSPHORIBOSYLFORMYLGLYCINAMIDINE SYNTHASE II (FGAM SYNTHASE II) [Bacillus subtilis]
1382	Bt1Gc1725	Bt1G3642	22451-21777	g4097532	772	787	3.10E-78	63	100	(U64311) phosphoribosylformylglycinamide synthetase I [Lactococcus lactis]
1382	Bt1Gc1725	Bt1G3643	23292-22698	g131631	668	671	6.00E-66	63	82	PHOSPHORIBOSYLAMINOIMIDAZOLE- SUCCINOCARBOXAMIDE SYNTHASE (SAICAR SYNTHETASE) (VEGETATIVE PROTEIN 286A) (VEG286A) [Bacillus subtilis]
1383	Bt1Gc1729	Bt1G3644	558-1	g1945721	598	532	3.20E-51	65	93	(Z94043) hypothetical protein [Bacillus subtilis]
1383	Bt1Gc1729	Bt1G3645	1679-576	g1945720	810	784	6.30E-78	46	100	(Z94043) hypothetical protein [Bacillus subtilis]
1383	Bt1Gc1729	Bt1G3646	2420-1689	g1945719	677	690	5.80E-68	49	100	(Z94043) hypothetical protein [Bacillus subtilis]
1383	Bt1Gc1729	Bt1G3647	3322-2423	g1945718	746	764	8.30E-76	49	100	(Z94043) hypothetical protein [Bacillus subtilis]
1383	Bt1Gc1729	Bt1G3648	3525-4268	g2443229	891	918	4.00E-92	69	100	(D86417) YfiG [Bacillus subtilis]
1383	Bt1Gc1729	Bt1G3649	11449-11692	g135254	124	149	7.80E-10	41	19	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3650	12010-12265	g135254	134	149	7.80E-10	39	21	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3651	12289-12550	g135254	170	189	3.30E-14	49	22	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3652	12289-12549	g135254	159	190	2.60E-14	49	21	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3653	12850-13123	g135254	226	263	1.70E-22	60	23	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3654	12850-13402	g135254	222	262	2.20E-22	64	20	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1383	Bt1Gc1729	Bt1G3655	12850-13963	g135254	220	270	2.60E-23	67	20	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3656	12850-14245	g135254	219	269	3.30E-23	77	16	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3657	12850-14524	g135254	221	277	3.60E-24	66	21	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3658	12850-14803	g135254	224	269	3.30E-23	55	26	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3659	12850-15085	g135254	221	260	3.80E-22	69	19	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3660	13129-15364	g135254	86	130	9.00E-08	54	11	TYPE III RESTRICTION-MODIFICATION SYSTEM ENZYME RES [Bacillus cereus]
1383	Bt1Gc1729	Bt1G3661	14022-15447	g732329	525	691	4.50E-68	35	100	HYPOTHETICAL 49.5 KD PROTEIN IN DAE-TYRZ INTERGENIC REGION PRECURSOR [Bacillus subtilis]
1383	Bt1Gc1729	Bt1G3662	16317-15469	g1789981	852	809	1.40E-80	54	100	(AE000433) IS150 putative transposase [Escherichia coli]
1384	Bt1Gc1730	Bt1G3663	1-484	g2635700	376	252	1.50E-21	47	81	(Z99120) similar to biotin metabolism [Bacillus subtilis]
1384	Bt1Gc1730	Bt1G3664	1420-545	g585225	1229	1161	7.10E-118	81	100	UTP--GLUCOSE-1-PHOSPHATE URIDYL TRANSFERASE (UDPGP) (ALPHA-D-PYROPHOSPHORYLASE) (UDPGP) [Bacillus subtilis]
1384	Bt1Gc1730	Bt1G3665	3252-1504	g2851553	1938	1958	2.50E-202	66	100	PROBABLE PHOSPHOMANNOMUTASE (PMM) [Bacillus subtilis]
1384	Bt1Gc1730	Bt1G3666	3393-3990	g3169326	341	389	4.60E-36	38	99	(AF026147) YojJ [Bacillus subtilis]
1384	Bt1Gc1730	Bt1G3667	5660-4164	g2635702	1481	1472	7.90E-151	58	100	(Z99120) similar to leucyl aminopeptidase [Bacillus subtilis]
1384	Bt1Gc1730	Bt1G3668	6446-5820	g2635704	533	505	2.30E-48	51	100	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1384	Bt1Gc1730	Bt1G3669	6853-6536	g2635705	310	196	1.30E-15	54	99	(Z99120) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1384	Bt1Gc1730	Bt1G3670	7455-6949	g2129134	116	229	4.10E-19	30	99	mutator protein mutT - Methanococcus jannaschii [Methanococcus jannaschii]
1384	Bt1Gc1730	Bt1G3671	8691-7480	g1934829	1256	1230	3.50E-125	62	100	(Z993939) unknown [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1384	Bt1Gc1730	Bt1G3672	9147-10142	g1934830	1319	1287	3.20E-131	74	100	(Z93939) thioredoxine reductase [Bacillus subtilis]
1384	Bt1Gc1730	Bt1G3673	26525-7739	g2133786	199	256	5.40E-20	26	64	NF-180 - sea lamprey [Petromyzon marinus]
1384	Bt1Gc1730	Bt1G3674	20695-21171	g2619047	271	323	4.50E-29	37	99	(AF027868) DNA-binding protein YobU [Bacillus subtilis]
1384	Bt1Gc1730	Bt1G3675	23563-24015	g1256140	410	415	8.00E-39	52	99	(AB002150) YbbK [Bacillus subtilis]
1385	Bt1Gc1731	Bt1G3676	761-267	g1922884	251	283	7.80E-25	36	99	(X99978) ORF7; hydrophobic protein [Lactobacillus plantarum]
1385	Bt1Gc1731	Bt1G3677	937-1809	g282295	436	509	8.80E-49	34	100	hypothetical protein, 35K - Lactococcus lactis [Lactococcus lactis]
1385	Bt1Gc1731	Bt1G3678	1876-2271	g1730902	330	336	1.90E-30	49	99	HYPOTHETICAL 14.7 KD PROTEIN IN BCSEA-DEGR INTERGENIC REGION [Bacillus subtilis]
1385	Bt1Gc1731	Bt1G3679	5089-4235	g2145400	409	428	3.40E-40	34	100	(Y09476) YitH [Bacillus subtilis]
1385	Bt1Gc1731	Bt1G3680	4975-24258	g4063042	124	199	3.50E-12	24	31	(AF068065) GP900; mucin-like glycoprotein [Cryptosporidium parvum]
1385	Bt1Gc1731	Bt1G3681	14392-19516	g2622173	321	649	2.00E-60	23	72	(AE000878) putative membrane protein [Methanobacterium thermoautotrophicum]
1385	Bt1Gc1731	Bt1G3682	24094-22005	g1730929	171	303	5.90E-27	25	90	HYPOTHETICAL 42.0 KD PROTEIN IN DAPB-PAPS INTERGENIC REGION [Bacillus subtilis]
1385	Bt1Gc1731	Bt1G3683	23836-23087	g3123077	387	405	9.20E-38	36	100	PUTATIVE GLYCOSYL TRANSFERASE HI0868 [Haemophilus influenzae Rd]
1386	Bt1Gc1732	Bt1G3684	1-685	g3005554	117	208	7.90E-17	20	72	(AF047044) putative transposase [Anabaena PCC7120]
1386	Bt1Gc1732	Bt1G3685	1157-1437	g2392827	79	123	7.00E-08	31	99	(AF011378) unknown [Bacteriophage sk1]
1386	Bt1Gc1732	Bt1G3686	2338-4026	g3128374	410	570	3.00E-55	28	100	(AF010496) hypothetical protein [Rhodobacter capsulatus]
1386	Bt1Gc1732	Bt1G3687	6002-5310	g3218350	169	240	2.80E-20	26	100	(AL023861) putative IS element transposase [Streptomyces coelicolor]
1386	Bt1Gc1732	Bt1G3688	7031-7444	g4126654	317	321	7.30E-29	39	99	(AB016282) ORF17 [bacteriophage phi-105]
1386	Bt1Gc1732	Bt1G3689	9373-8583	g732244	223	323	4.50E-29	31	100	HYPOTHETICAL 24.0 KD PROTEIN IN PONA-COTD INTERGENIC REGION (ORF1) [Bacillus subtilis]
1386	Bt1Gc1732	Bt1G3690	11248-10047	g1652146	122	271	1.50E-23	24	84	(D90903) hypothetical protein [Synecocystis sp.]
1386	Bt1Gc1732	Bt1G3691	17140-18847	g928836	120	206	1.10E-16	34	51	(L44593) repressor protein [Lactococcus lactis phase BK5-T]
1386	Bt1Gc1732	Bt1G3692	19441-21005	g1865708	521	608	2.80E-59	35	100	(Y11478) probably site-specific recombinase of the resolvase family of enzymes [Bacteriophage TP21]
1386	Bt1Gc1732	Bt1G3693	22446-20601	g1934835	2207	2243	1.60E-232	66	100	(Z93940) asparagine synthetase [Bacillus subtilis]
1387	Bt1Gc1733	Bt1G3694	2648-329	g3128293	109	244	3.50E-20	23	75	(AF010496) hypothetical protein [Rhodobacter capsulatus]
1387	Bt1Gc1733	Bt1G3695	3003-1256	g113673	694	648	1.60E-63	79	70	N-ACETYL MURAMOYL-L-ALANINE AMIDASE

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1387	Bt1Gc1733	Bt1G3696	3428-3036	g141088	180	189	7.10E-15	32	99	CWLA PRECURSOR (CELL WALL HYDROLASE) (AUTOLYSIN) [Bacillus sp.]
1387	Bt1Gc1733	Bt1G3697	3770-3473	g2160199	97	133	6.10E-09	34	95	HYPOTHETICAL 14.9 KD PROTEIN IN NAGH 3'REGION (ORFD) [Clostridium perfringens]
1387	Bt1Gc1733	Bt1G3698	12383-10534	g4126623	194	231	2.50E-19	31	79	(D13377) XpaF1 protein [Bacillus licheniformis]
1387	Bt1Gc1733	Bt1G3699	16615-12628	g3947462	240	364	1.40E-32	27	57	(AB016282) ORF37 [bacteriophage phi-105]
1387	Bt1Gc1733	Bt1G3700	16154-13023	g4249699	75	193	5.10E-13	19	41	(AJ006589) gp43 [Bacteriophage phi-C31]
1387	Bt1Gc1733	Bt1G3701	19546-18938	g320252	258	326	2.20E-29	34	99	(AF097906) myosin heavy chain [Rana catesbeiana]
1387	Bt1Gc1733	Bt1G3702	20104-19580	g320251	126	180	6.40E-14	42	99	promoter inhibitor protein BpL - Lactococcus lactis phage BK5-T []
1387	Bt1Gc1733	Bt1G3703	20831-19158	g4530143	166	236	7.50E-20	27	88	hypothetical protein 1 (bpi 5' region) - Lactococcus lactis phage BK5-T (fragment) []
1388	Bt1Gc1739	Bt1G3704	1-353	g3005554	97	135	1.50E-08	24	36	(AF085222) putative scaffolding protein [Streptococcus thermophilus bacteriophage DT1]
1388	Bt1Gc1739	Bt1G3705	4302-2520	g2444121	62	138	4.40E-08	42	6	(AF047044) putative transposase [Anabaena PCC7120]
1388	Bt1Gc1739	Bt1G3706	8733-6448	g2764873	342	411	5.60E-38	30	47	(U88974) ORF42 [Streptococcus thermophilus bacteriophage O1205]
1388	Bt1Gc1739	Bt1G3707	10634-7189	g3282276	307	440	1.40E-61	25	100	(X97918) gene 18.1 [Bacteriophage SPP1]
1388	Bt1Gc1739	Bt1G3708	16049-14381	g4530143	251	257	4.40E-22	31	94	(AF009630) I16 [bacteriophage bIL170]
1388	Bt1Gc1739	Bt1G3709	20871-21284	g4126654	319	323	4.50E-29	39	99	(AF085222) putative scaffolding protein [Streptococcus thermophilus bacteriophage DT1]
1388	Bt1Gc1739	Bt1G3710	23214-22413	g732244	227	329	1.00E-29	32	100	(AB016282) ORF17 [bacteriophage phi-105]
1388	Bt1Gc1739	Bt1G3711	26292-23643	g1652146	116	266	4.90E-23	22	84	HYPOTHETICAL 24.0 KD PROTEIN IN PONA-COTD INTERGENIC REGION (ORF1) [Bacillus subtilis]
1389	Bt1Gc1734	Bt1G3712	958-524	g2634157	495	506	1.80E-48	63	99	(D90903) hypothetical protein [Synechocystis sp.]
1389	Bt1Gc1734	Bt1G3713	2418-2876	g1881322	412	348	1.00E-31	55	99	(Z99113) yndB [Bacillus subtilis]
1389	Bt1Gc1734	Bt1G3714	3755-2973	g729328	1198	1198	8.50E-122	88	100	(AB001488) SIMILAR TO TRANSCRIPTION FACTOR OF MYXOCOCCUS XANTHUS. [Bacillus subtilis]
1389	Bt1Gc1734	Bt1G3715	4626-3772	g731355	1227	1100	2.10E-111	82	100	GLUCOSE 1-DEHYDROGENASE [Bacillus megaterium]
1389	Bt1Gc1734	Bt1G3716	5870-5405	g2633801	505	514	2.60E-49	61	99	HYPOTHETICAL 30.5 KD PROTEIN IN GDHI 5'REGION (ORF 2) [Bacillus megaterium]
1389	Bt1Gc1734	Bt1G3717	7647-6358	g2633799	1392	1334	3.30E-136	63	100	(Z99111) molybdopterin converting factor (subunit 2) [Bacillus subtilis]
1389	Bt1Gc1734	Bt1G3717	7647-6358	g2633799	1392	1334	3.30E-136	63	100	(Z99111) molybdopterin biosynthesis protein [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1389	Bt1Gc1734	Bt1G3718	7731-8248	g1945112	583	556	9.20E-54	68	99	(D88802) E. coli moaC protein; P30747 (368) [Bacillus subtilis]
1389	Bt1Gc1734	Bt1G3719	9267-8257	g2633524	919	909	3.60E-91	55	100	(Z99110) similar to thiamin biosynthesis [Bacillus subtilis]
1389	Bt1Gc1734	Bt1G3720	10306-9287	g730103	1008	962	8.70E-97	57	100	NARA PROTEIN [Bacillus subtilis]
1389	Bt1Gc1734	Bt1G3721	15462-15259	g2635515	243	255	7.20E-22	66	99	(Z99119) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1389	Bt1Gc1734	Bt1G3722	17124-16435	g3290174	1214	1214	1.70E-123	99	100	(AF067645) homoserine O-acetyltransferase homolog [Bacillus cereus]
1389	Bt1Gc1734	Bt1G3723	17276-19498	g3290175	3186	2273	1.00E-237	95	100	(AF067645) spore germination protein GerIA [Bacillus cereus]
1389	Bt1Gc1734	Bt1G3724	19518-20609	g3290176	1738	1637	2.60E-168	92	100	(AF067645) spore germination protein GerIB [Bacillus cereus]
1389	Bt1Gc1734	Bt1G3725	25860-24146	g3123286	2411	2330	9.50E-242	77	71	LEUCYL-TRNA SYNTHETASE (LEUCINE--TRNA LIGASE) (LEURS) [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3726	597-1	g1075196	578	521	9.20E-50	55	26	pyruvate formate-lyase (pfI) homolog - Haemophilus influenzae (strain Rd KW20) [Haemophilus influenzae Rd]
1390	Bt1Gc1736	Bt1G3727	4722-2534	g3256832	776	841	5.80E-84	41	100	(AP000002) 455aa long hypothetical Vi polysaccharide biosynthesis protein [Pyrococcus horikoshii]
1390	Bt1Gc1736	Bt1G3728	8237-6999	g2634067	825	857	1.20E-85	48	100	(Z99112) penicillin-binding protein [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3729	9993-8491	g2443228	1358	1186	1.60E-120	61	100	(D86417) Yfif [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3730	10319-11245	g2494800	928	928	3.50E-93	56	100	HYPOTHETICAL 34.0 KD PROTEIN IN NPPE-PYCA INTERGENIC REGION [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3731	14829-13552	g2633908	1485	1492	6.00E-153	63	100	(Z99112) similar to acetylornithine deacetylase [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3732	16355-14952	g730601	1510	1400	3.40E-143	58	100	AMINO-ACID PERMEASE ROCE [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3733	17016-18407	g1064808	1348	1394	1.50E-142	57	100	(D78193) positive regulatory protein [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3734	19530-18796	g3915176	813	773	9.30E-77	62	100	TRNA PSEUDOURIDINE SYNTHASE A (PSEUDOURIDYLATE SYNTHASE I) (PSEUDOURIDINE SYNTHASE I) (URACIL HYDROLYASE) [Bacillus sp.]
1390	Bt1Gc1736	Bt1G3735	20972-19581	g2633126	1632	1537	1.00E-157	66	100	(Z99108) similar to RNA methyltransferase [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3736	21840-20980	g4584142	1431	1431	1.70E-146	93	100	(AJ010139) DNA alkylation repair enzyme [Bacillus cereus]
1390	Bt1Gc1736	Bt1G3737	22682-21903	g4584141	1375	1311	9.10E-134	99	100	(AJ010139) polysaccharide deacetylase-like protein

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aaf_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1390	Bt1Gc1736	Bt1G3738	24330-22813	g417009	2132	2155	3.30E-223	77	100	[Bacillus cereus] FUMARATE HYDRATASE CLASS I, AEROBIC (FUMARASE) [Bacillus stearothermophilus]
1390	Bt1Gc1736	Bt1G3739	24929-25390	g2626813	680	680	6.70E-67	80	99	(D83967) YfkA [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3740	25589-26047	g2626814	576	578	4.30E-56	71	99	(D83967) YfkB [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3741	26892-26095	g2626816	721	725	1.10E-71	54	100	(D83967) YfkD [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3742	28021-26972	g2626817	1306	1010	7.10E-102	73	100	(D83967) YfkE [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3743	28194-28616	g2829682	285	294	5.30E-26	38	99	GENERAL STRESS PROTEIN 26 (GSP26) [Bacillus subtilis]
1390	Bt1Gc1736	Bt1G3744	30162-29677	g2632514	340	262	1.30E-22	41	99	(Z99105) similar to alkaline phosphatase [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3745	6139-5237	g1749770	735	735	9.90E-73	51	100	(Y09946) transposase [Bacillus thuringiensis]
1391	Bt1Gc1738	Bt1G3746	6570-7908	g4103625	1050	927	4.50E-93	49	100	(AF026470) gluconate permease [Pseudomonas aeruginosa]
1391	Bt1Gc1738	Bt1G3747	11833-10575	g709994	598	586	6.10E-57	39	100	(D14399) hypothetical protein [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3748	12582-11811	g1176995	835	851	5.00E-85	61	100	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN IN IDH 3'REGION [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3749	16606-12618	g1731300	857	757	4.60E-75	36	100	HYPOTHETICAL 70.5 KD PROTEIN IN IDH 3'REGION [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3750	17723-16743	g1176994	566	631	1.00E-61	38	100	HYPOTHETICAL SENSOR-LIKE HISTIDINE KINASE IN IDH 3'REGION [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3751	18421-17735	g1176993	657	661	6.90E-65	55	100	HYPOTHETICAL 26.6 KD SENSORY TRANSDUCTION PROTEIN IN IDH 3'REGION [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3752	18631-18900	g2635892	333	336	1.90E-30	67	99	(Z99121) similar to transcriptional regulator (ArsR family) [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3753	21117-20248	g1881325	1211	1211	3.60E-123	77	100	(AB001488) PROBABLE HTH_ARAC_FAMILY OF TRANSCRIPTIONAL REGULATOR. [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3754	21427-22185	g2293177	707	712	2.70E-70	54	100	(AF008220) transporter [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3755	29515-28628	g728972	582	602	1.20E-58	39	100	BMRU PROTEIN [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3756	30230-29691	g585264	447	460	1.40E-43	46	99	HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE (HGPRT) (HGPRTASE) [Bacillus subtilis]
1391	Bt1Gc1738	Bt1G3757	32538-33239	g1575577	400	390	3.60E-36	39	100	(U67196) DNA-binding response regulator [Thermotoga maritima]
1391	Bt1Gc1738	Bt1G3758	32736-35074	g466195	306	372	7.30E-34	35	41	SENSOR PROTEIN RESE [Bacillus subtilis]
1392	Bt1Gc1735	Bt1G3759	3545-497	g4062401	1252	1414	1.10E-144	47	93	(D90720) Hypothetical ABC transporter ATP-binding protein in bcr 5'region. [Escherichia coli]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1392	Bt1Gc1735	Bt1G3760	3362-2421	g129184	538	462	8.40E-44	39	100	OLIGOPEPTIDE TRANSPORT SYSTEM PERMEASE PROTEIN OPPC [Bacillus subtilis]
1392	Bt1Gc1735	Bt1G3761	3714-3348	g3757815	285	337	1.50E-30	46	38	(AF042861) putative OppB [Treponema denticola]
1392	Bt1Gc1735	Bt1G3762	7255-3852	g136144	2443	2424	1.00E-251	100	100	TRANSPORSEASE FOR INSERTION SEQUENCE ELEMENT IS231C [Bacillus thuringiensis]
1392	Bt1Gc1735	Bt1G3763	8072-8539	g2635598	242	270	1.90E-23	34	99	(Z99119) similar to cysteine dioxygenase [Bacillus subtilis]
1392	Bt1Gc1735	Bt1G3764	8846-9980	g2618997	531	545	1.30E-52	33	100	(AF027868) transcription antiterminator [Bacillus subtilis]
1392	Bt1Gc1735	Bt1G3765	11562-10234	g731354	1290	1306	3.10E-133	55	100	HYPOTHETICAL 50.8 KD PROTEIN IN SRFA4-SFP INTERGENIC REGION (ORF8) [Bacillus subtilis]
1393	Bt1Gc1737	Bt1G3766	1656-1	g3915204	2286	2287	3.40E-237	79	58	EXCINUCLEASE ABC SUBUNIT A [Bacillus subtilis]
1393	Bt1Gc1737	Bt1G3767	3644-1671	g3123297	2821	2499	1.20E-259	83	100	EXCINUCLEASE ABC SUBUNIT B (DINA PROTEIN) [Bacillus subtilis]
1393	Bt1Gc1737	Bt1G3768	7813-8709	g732377	642	649	1.30E-63	45	100	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN MMR-PTA INTERGENIC REGION [Bacillus subtilis]
1393	Bt1Gc1737	Bt1G3769	9943-11574	g2769708	955	852	3.90E-85	37	100	(U82085) pristinamycin resistance protein VgaB [Staphylococcus aureus]
1393	Bt1Gc1737	Bt1G3770	13770-12046	g1176951	2189	2199	7.20E-228	72	100	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN IN ACDA 5'REGION [Bacillus subtilis]
1393	Bt1Gc1737	Bt1G3771	15228-14038	g2618836	654	624	5.70E-61	36	100	(AF017113) YvjD [Bacillus subtilis]
1393	Bt1Gc1737	Bt1G3772	16771-15335	g2618834	1325	1349	8.50E-138	53	100	(AF017113) putative protease [Bacillus subtilis]
1393	Bt1Gc1737	Bt1G3773	17780-16893	g2618835	898	701	4.00E-69	59	100	(AF017113) cell division protein [Bacillus subtilis]
1393	Bt1Gc1737	Bt1G3774	18459-17776	g2618833	752	757	4.60E-75	64	100	(AF017113) cell division ATP-binding protein [Bacillus subtilis]
1393	Bt1Gc1737	Bt1G3775	20607-19509	g2331287	1560	1491	7.60E-153	81	100	(AF013188) release factor 2 [Bacillus subtilis]
1393	Bt1Gc1737	Bt1G3776	21742-20755	g3122850	1195	1169	1.00E-118	68	70	PREPROTEIN TRANSLOCASE SECA SUBUNIT [Bacillus firmus]
1394	Bt1Gc1758	Bt1G3777	1-282	g1724002	139	159	2.90E-11	44	30	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN 2 IN GLPD-CSPB INTERGENIC REGION [Bacillus subtilis]
1395	Bt1Gc1762	Bt1G3778	212-1	g3290175	244	265	5.70E-22	78	11	(AF067645) spore germination protein GerIA [Bacillus cereus]
1396	Bt1Gc1765	Bt1G3779	1-222	g3123300	191	208	7.10E-17	56	23	THIOREDOXIN REDUCTASE (GENERAL STRESS PROTEIN 35) [Bacillus subtilis]
1397	Bt1Gc1763	Bt1G3780	270-117	g595780	122	141	8.70E-10	60	27	(U13871) lacZ alpha peptide [Cloning vector]
1398	Bt1Gc1767	Bt1G3781	47-478	g1934652	478	397	6.50E-37	65	33	(U93876) amino acid transporter [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1399	Bt1Gc1804	Bt1G3783	122-1	g3128280	94	110	1.00E-05	49	11	(AF010496) hypothetical ABC transporter A TP-binding protein [Rhodobacter capsulatus]
1400	Bt1Gc1822	Bt1G3784	13-231	g321919	102	97	4.00E-05	38	50	hypothetical 16.9K protein - Salmonella typhimurium plasmid NTP16 []
1401	Bt1Gc1818	Bt1G3785	239-1	g4584140	234	256	5.70E-22	66	29	(AJ010139) fumarate hydratase [Bacillus cereus]
1402	Bt1Gc1828	Bt1G3787	176-1	g4584140	116	150	1.90E-10	53	21	(AJ010139) fumarate hydratase [Bacillus cereus]
1403	Bt1Gc1832	Bt1G3788	137-1	g124464	135	156	3.20E-10	60	7	IMMUNE INHIBITOR A PRECURSOR [Bacillus thuringiensis]
1404	Bt1Gc1771	Bt1G3789	1-276	g1172949	277	280	1.60E-24	66	55	50S RIBOSOMAL PROTEIN L10 (BL5) (COLD ACCLIMATIZATION PROTEIN) (CAP) (VEGETATIVE PROTEIN 300) (VEG300) [Bacillus subtilis]
1405	Bt1Gc1845	Bt1G3790	416-1	g4323583	73	113	7.70E-06	21	30	(AF101076) maturase-related protein [Pseudomonas putida]
1406	Bt1Gc1849	Bt1G3791	281-1	g2462090	154	181	3.60E-13	39	19	(Y11171) BC542A protein [Bacillus cereus]
1407	Bt1Gc1862	Bt1G3792	1-388	g1172446	98	144	6.90E-09	26	18	KETO-ACID FORMATE ACETYLTRANSFERASE (KETO-ACID FORMATE-LYASE) [Escherichia coli]
1408	Bt1Gc1869	Bt1G3794	41-488	g3401958	562	565	1.00E-54	74	82	Pyrr, The Bacillus Subtilis Pyrimidine Biosynthetic Operon Repressor, Dimeric Form [Bacillus subtilis]
1409	Bt1Gc1866	Bt1G3795	267-1	g143324	203	159	5.50E-11	54	23	(M37169) APase I [Bacillus licheniformis]
1410	Bt1Gc1876	Bt1G3796	397-1	g2619017	92	151	8.00E-10	33	23	(AF027868) squalene-hopene cyclase [Bacillus subtilis]
1411	Bt1Gc1891	Bt1G3798	534-1	g2501426	691	706	1.20E-69	80	41	PYRIMIDINE-NUCLEOSIDE PHOSPHORYLASE [Bacillus stearothermophilus]
1412	Bt1Gc1896	Bt1G3799	274-62	g2984723	190	194	2.10E-15	58	99	(AF053927) probable spore germination protein F [Bacillus cereus]
1413	Bt1Gc1898	Bt1G3800	1-235	g586902	132	153	2.10E-10	47	22	HYPOTHETICAL 40.7 KD PROTEIN IN MECB-GLTX INTERGENIC REGION [Bacillus subtilis]
1414	Bt1Gc1894	Bt1G3801	61-299	g2116759	124	140	9.10E-09	36	17	(D86418) YfnA [Bacillus subtilis]
1415	Bt1Gc1917	Bt1G3802	169-1	g3290175	80	114	9.60E-06	46	8	(AF067645) spore germination protein GerIA [Bacillus cereus]
1416	Bt1Gc1960	Bt1G3803	1-508	g2635490	462	484	3.90E-46	59	53	(Z99119) similar to hypothetical proteins [Bacillus subtilis]
1417	Bt1Gc1981	Bt1G3804	1-224	g2117766	65	120	1.00E-06	42	18	orthinine aminotransferase rocD - Bacillus subtilis []
1418	Bt1Gc1983	Bt1G3805	515-1	g3401988	242	293	2.10E-25	38	34	Perfringolysin O []
1419	Bt1Gc1979	Bt1G3806	1-436	g118797	152	170	3.20E-12	33	38	DNA POLYMERASE III, BETA CHAIN [Bacillus subtilis]
1420	Bt1Gc1989	Bt1G3807	117-323	g4584092	172	181	5.00E-14	54	99	(AJ010129) hypothetical protein [Bacillus cereus]
1421	Bt1Gc1991	Bt1G3808	1-381	g2983101	128	183	3.10E-13	32	22	(AF000690) acetolactate synthase large subunit [Aquifex aeolicus]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1422	Bt1Gc1992	Bt1G3809	309-1	g1708465	220	250	1.60E-20	50	18	DIHYDROXY-ACID DEHYDRATASE (DAD) (VEGETATIVE PROTEIN 110) (VEG110) [Bacillus subtilis]
1423	Bt1Gc2025	Bt1G3810	1-167	g144312	114	123	7.00E-08	53	46	(J01566) 13.8 kd ORF [Plasmid ColE1]
1424	Bt1Gc2027	Bt1G3811	1-491	g134774	181	207	2.30E-16	42	44	STAGE V SPORULATION PROTEIN E [Bacillus subtilis]
1425	Bt1Gc2028	Bt1G3812	423-117	g140583	204	220	3.70E-18	46	33	UDP-N-ACETYLENOLPYRUVOYLGLUCOSAMINE REDUCTASE (UDP-N-ACETYLMURAMATE DEHYDROGENASE) [Bacillus subtilis]
1426	Bt1Gc2038	Bt1G3813	1-212	g2827439	122	143	3.70E-09	49	17	(AF043609) aluminum resistance protein [Arthrobacter viscosus]
1427	Bt1Gc2029	Bt1G3814	397-1	g1945649	168	191	4.40E-15	35	84	(Z94043) hypothetical protein [Bacillus subtilis]
1428	Bt1Gc2057	Bt1G3815	186-1	g118516	79	109	1.40E-05	37	16	ALANINE DEHYDROGENASE [Bacillus stearothermophilus]
1429	Bt1Gc2060	Bt1G3816	249-1	g1881268	136	132	7.80E-08	53	16	(AB001488) ATP-DEPENDENT RNA HELICASE DEAD HOMOLOG. [Bacillus subtilis]
1430	Bt1Gc2067	Bt1G3817	255-1	g1001657	181	184	3.30E-14	42	28	(D64002) acyl- [Synecocystis sp.]
1431	Bt1Gc2077	Bt1G3818	160-507	g2632987	212	262	4.10E-22	44	25	(Z99107) alternate gene name: yerS; similar to RNA methyltransferase [Bacillus subtilis]
1432	Bt1Gc2081	Bt1G3819	1-195	g1787795	99	126	1.70E-07	38	18	(AE000249) putative LACI-type transcriptional regulator [Escherichia coli]
1433	Bt1Gc2102	Bt1G3820	1-214	g4699725	140	163	3.20E-11	45	14	Bacillus Licheniformis Alpha-Amylase []
1434	Bt1Gc2111	Bt1G3822	1-352	g2633515	155	178	1.00E-13	36	44	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
1435	Bt1Gc2117	Bt1G3823	243-1	g3914289	126	159	1.40E-10	44	12	TOPOISOMERASE IV SUBUNIT B [Bacillus subtilis]
1436	Bt1Gc2110	Bt1G3824	1-423	g1220104	162	179	2.30E-13	31	41	(D49784) sodium-coupled branched-chain amino acid carrier [Clostridium perfringens]
1437	Bt1Gc2120	Bt1G3825	371-1	g1730918	419	433	9.90E-41	71	85	HYPOTHETICAL 15.9 KD PROTEIN IN ILVD-THYB INTERGENIC REGION [Bacillus subtilis]
1438	Bt1Gc2122	Bt1G3826	270-1	g118515	148	165	1.10E-11	52	23	ALANINE DEHYDROGENASE [Bacillus sphaericus]
1439	Bt1Gc2134	Bt1G3827	340-1	g1934652	197	208	3.30E-16	40	26	(U93876) amino acid transporter [Bacillus subtilis]
1440	Bt1Gc2138	Bt1G3828	316-1	g586880	171	189	7.10E-15	36	37	PUR OPERON REPRESSOR [Bacillus subtilis]
1441	Bt1Gc2158	Bt1G3829	328-1	g2612901	188	222	6.20E-18	52	28	(AF015825) NADH dehydrogenase-like protein [Bacillus subtilis]
1442	Bt1Gc2162	Bt1G3830	1-280	g473955	105	156	1.40E-10	40	22	(D29674) DNA-binding protein [Lactobacillus sp.]
1443	Bt1Gc2141	Bt1G3831	441-220	g1001739	104	119	4.60E-07	32	30	(D64004) ABC transporter [Synecocystis sp.]
1444	Bt1Gc2171	Bt1G3832	541-1	g3122117	222	212	3.10E-16	41	27	CELL DIVISION PROTEIN FTSH HOMOLOG [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1445	Bt1Gc2172	Bt1G3833	248-1	g225559	220	263	3.70E-22	60	17	ORF IS231C [Bacillus thuringiensis]
1446	Bt1Gc2210	Bt1G3835	1-542	g4514362	398	360	5.40E-33	45	57	(AB013378) YhaQ [Bacillus halodurans]
1447	Bt1Gc2236	Bt1G3836	256-1	g482286	185	207	3.10E-15	54	6	DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain (version 2) - Bacillus subtilis []
1448	Bt1Gc2242	Bt1G3837	1-251	g3915460	256	285	4.80E-25	61	37	PUTATIVE NAD(P)H NITROREDUCTASE YFKO [Bacillus subtilis]
1449	Bt1Gc2247	Bt1G3838	1-388	g585920	298	323	1.00E-27	57	11	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Bacillus subtilis]
1450	Bt1Gc2251	Bt1G3839	168-1	g1881236	118	123	1.20E-06	46	8	(AB001488) PROBABLE DNA TOPOISOMERASE III [Bacillus subtilis]
1451	Bt1Gc2249	Bt1G3840	228-1	g285621	108	110	8.60E-06	38	24	(D13095) undefined open reading frame [Bacillus stearothermophilus]
1452	Bt1Gc2264	Bt1G3841	305-101	g2116756	147	167	1.80E-11	49	11	(D86418) YfmR [Bacillus subtilis]
1453	Bt1Gc2266	Bt1G3842	161-278	g2636203	127	130	1.30E-08	57	51	(Z99122) ywzB [Bacillus subtilis]
1454	Bt1Gc2252	Bt1G3843	366-1	g2635801	146	190	3.40E-14	37	26	(Z99120) fumarate hydratase [Bacillus subtilis]
1455	Bt1Gc2284	Bt1G3844	1-419	g2454562	136	61	0.014	32	55	(AF015609) unknown [Bacillus subtilis]
1456	Bt1Gc2298	Bt1G3845	248-1	g1770025	87	115	4.30E-06	33	18	(Z75208) hypothetical protein [Bacillus subtilis]
1457	Bt1Gc2294	Bt1G3846	1-229	g207999	261	274	7.00E-24	71	78	(M38482) HBSu protein [Artificial gene]
1458	Bt1Gc2301	Bt1G3847	53-419	g1652657	75	137	2.20E-08	34	17	(D90907) hypothetical protein [Synechocystis sp.]
1459	Bt1Gc2302	Bt1G3848	290-1	g3955201	106	125	2.10E-07	34	29	(AF022796) MoeB [Staphylococcus carnosus]
1460	Bt1Gc2300	Bt1G3849	158-1	g994737	104	125	1.30E-07	52	20	(M18327) kanamycin resistance protein [cloning vectors]
1461	Bt1Gc2305	Bt1G3850	1-315	g847865	70	164	3.20E-12	39	88	(U26464) beta-galactosidase alpha [Cloning vector pZC320]
1462	Bt1Gc2311	Bt1G3851	1-328	g1934835	114	149	1.50E-09	32	18	(Z93940) asparagine synthetase [Bacillus subtilis]
1463	Bt1Gc2330	Bt1G3852	1-252	g2635183	97	153	2.10E-10	36	23	(Z99117) similar to folate metabolism [Bacillus subtilis]
1464	Bt1Gc2337	Bt1G3853	1-389	g225559	184	239	1.70E-19	42	27	ORF IS231C [Bacillus thuringiensis]
1465	Bt1Gc2340	Bt1G3854	285-1	g1894750	215	241	2.20E-20	49	39	(Z92952) product similar to E.coli YjaF protein [Bacillus subtilis]
1466	Bt1Gc2368	Bt1G3855	107-228	g2226150	104	115	5.00E-07	54	24	(Y14080) hypothetical protein [Bacillus subtilis]
1467	Bt1Gc2378	Bt1G3856	1-237	g3123227	216	234	7.10E-19	50	15	GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1468	Bt1Gc2402	Bt1G3857	1-341	g1770035	180	217	6.30E-17	36	20	(Z75208) hypothetical protein [Bacillus subtilis]
1469	Bt1Gc2406	Bt1G3858	120-1	g586617	166	169	9.40E-13	83	35	HYPOTHETICAL 13.2 KD PROTEIN IN FFH 5'REGION [Bacillus subtilis]
1469	Bt1Gc2406	Bt1G3859	534-254	g2633967	298	327	1.70E-29	69	28	(Z99112) signal recognition particle (docking protein) [Bacillus subtilis]
1470	Bt1Gc2410	Bt1G3860	389-1	g3121784	364	358	5.30E-32	59	19	POTASSIUM-TRANSPORTING ATPASE B CHAIN (ATP PHOSPHOHYDROLASE [Clostridium acetobutylicum])
1471	Bt1Gc2412	Bt1G3861	558-1	g2633162	161	156	1.40E-10	25	44	(Z99108) similar to multidrug resistance protein [Bacillus subtilis]
1472	Bt1Gc2413	Bt1G3862	178-1	g4512347	101	125	4.00E-07	47	12	(AB011836) similar to B.subtilis yerN gene(87% identity) [Bacillus halodurans]
1473	Bt1Gc2425	Bt1G3863	235-1	g2851477	171	207	7.30E-16	52	14	ARGINYL-TRNA SYNTHETASE (ARGININE--TRNA LIGASE) (ARGRS) [Bacillus subtilis]
1474	Bt1Gc2428	Bt1G3864	1-407	g2495457	222	254	1.40E-21	43	34	HYPOTHETICAL 41.5 KD PROTEIN IN AMHX-AMYE INTERGENIC REGION [Bacillus subtilis]
1475	Bt1Gc2431	Bt1G3865	309-1	g1731001	161	187	1.20E-14	42	37	HYPOTHETICAL 30.3 KD PROTEIN IN GLYS-DNAG/DNAE INTERGENIC REGION [Bacillus subtilis]
1476	Bt1Gc2465	Bt1G3866	1-188	g3121784	132	105	9.20E-05	51	9	POTASSIUM-TRANSPORTING ATPASE B CHAIN (ATP PHOSPHOHYDROLASE [Clostridium acetobutylicum])
1477	Bt1Gc2467	Bt1G3867	156-371	g1162915	179	224	1.40E-18	60	28	(L48553) virA gene product [Bacillus anthracis]
1478	Bt1Gc2471	Bt1G3868	66-533	g4512410	605	608	2.80E-59	77	71	(AB017508) rpsC homologue (identity of 87% to B. subtilis) [Bacillus halodurans]
1479	Bt1Gc2472	Bt1G3869	377-1	g132987	359	391	2.80E-36	64	69	50S RIBOSOMAL PROTEIN L5 []
1480	Bt1Gc2480	Bt1G3870	461-1	g1934809	135	186	1.20E-13	33	29	(Z93937) unknown [Bacillus subtilis]
1481	Bt1Gc2482	Bt1G3871	1-224	g1703223	125	153	2.80E-10	45	18	ASPARTOKINASE 2 ALPHA AND BETA SUBUNITS (ASPARTATE KINASE 2) [Bacillus stearothermophilus]
1482	Bt1Gc2497	Bt1G3872	367-1	g2635819	190	226	1.60E-18	39	33	(Z99120) similar to two-component response regulator [Bacillus subtilis]
1483	Bt1Gc2506	Bt1G3873	457-1	g2649148	194	239	3.60E-20	35	68	(AE001005) aspartate racemase [Archaeoglobus fulgidus]
1484	Bt1Gc2510	Bt1G3874	1-187	g4512396	125	146	7.60E-09	54	5	(AB017508) rpoB homologue (identity of 87% to B. subtilis) [Bacillus halodurans]
1485	Bt1Gc2511	Bt1G3875	402-74	g1064795	196	242	1.70E-20	50	68	(D78193) yydA [Bacillus subtilis]
1486	Bt1Gc2507	Bt1G3876	1-288	g135106	190	208	4.20E-16	48	19	GLUTAMYL-TRNA SYNTHETASE (GLUTAMATE--TRNA LIGASE) (GLURS) [Bacillus subtilis]
1487	Bt1Gc2535	Bt1G3877	1-117	g1944414	89	110	3.30E-05	57	5	(D87026) glycogen phosphorylase [Bacillus

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aaf_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1488	Bt1Gc2532	Bt1G3878	20-387	g221919	152	166	2.00E-12	38	83	stearothermophilus] hypothetical 16.9K protein - Salmonella typhimurium plasmid NTP16 []
1489	Bt1Gc2536	Bt1G3879	1-221	g1724002	129	139	4.90E-09	44	25	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN 2 IN GLPD-CSPB INTERGENIC REGION [Bacillus subtilis]
1490	Bt1Gc2559	Bt1G3880	1-351	g2500058	259	252	1.80E-20	54	16	FORMATE ACETYLTRANSFERASE (PYRUVATE FORMATE-LYASE) [Clostridium pasteurianum]
1491	Bt1Gc2560	Bt1G3881	457-1	g1730908	210	139	8.20E-09	35	40	HYPOTHETICAL 43.6 KD PROTEIN IN CP5D-METB INTERGENIC REGION [Bacillus subtilis]
1492	Bt1Gc2563	Bt1G3882	455-1	g2635675	428	386	9.50E-36	56	41	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1493	Bt1Gc2567	Bt1G3883	388-1	g2293298	173	213	2.00E-17	41	56	(AF008220) putative transcription regulator [Bacillus subtilis]
1494	Bt1Gc2577	Bt1G3885	1-298	g130130	162	134	5.70E-08	43	17	ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN PHOR [Bacillus subtilis]
1495	Bt1Gc2578	Bt1G3886	1-267	g1730943	131	149	1.20E-10	39	60	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN UVRX-JLVA INTERGENIC REGION [Bacillus subtilis]
1496	Bt1Gc2570	Bt1G3887	1-289	g132246	105	74	0.085	31	26	RECF PROTEIN [Bacillus subtilis]
1497	Bt1Gc2593	Bt1G3888	1-234	g3821797	88	139	1.40E-08	46	14	(D10594) chitinase D precursor [Bacillus circulans]
1498	Bt1Gc2594	Bt1G3889	1-355	g1709733	79	152	5.80E-10	37	15	PHOSPHOENOLPYRUVATE CARBOXYKINASE (ATP) [Staphylococcus aureus]
1499	Bt1Gc2608	Bt1G3890	1-387	g2635780	164	229	6.20E-18	38	16	(Z99120) similar to 3-hydroxyacyl-CoA dehydrogenase [Bacillus subtilis]
1500	Bt1Gc2613	Bt1G3891	390-1	g2634117	137	149	4.20E-10	32	41	(Z99113) tRNA isopentenylpyrophosphate transferase [Bacillus subtilis]
1501	Bt1Gc2617	Bt1G3892	1-317	g1205984	180	127	5.60E-07	47	12	(U33536) DNA polymerase I [Bacillus stearothermophilus]
1502	Bt1Gc2622	Bt1G3893	1-441	g2337795	270	217	1.40E-16	48	16	(Y13937) putative PacL protein [Bacillus subtilis]
1503	Bt1Gc2629	Bt1G3894	1-245	g4584142	147	190	5.60E-15	50	28	(AJ010139) DNA alkylation repair enzyme [Bacillus cereus]
1504	Bt1Gc2631	Bt1G3895	383-1	g1730999	247	305	3.20E-26	49	17	HYPOTHETICAL 79.2 KD PROTEIN IN PHOH-DGKA INTERGENIC REGION [Bacillus subtilis]
1505	Bt1Gc2632	Bt1G3896	1-322	g133732	431	441	1.40E-41	79	76	30S RIBOSOMAL PROTEIN S12 (BS12) []
1506	Bt1Gc2639	Bt1G3897	1-403	g1072499	251	214	3.50E-17	44	37	phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) - Bacillus subtilis [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1507	Bt1Gc2641	Bt1G3898	73-413	g121466	419	352	3.80E-32	72	27	PROTON/SODIUM-GLUTAMATE SYMPORT PROTEIN (GLUTAMATE-ASPARTATE CARRIER PROTEIN) [Bacillus caldotenax]
1508	Bt1Gc2647	Bt1G3899	330-1	g1170998	132	144	4.20E-10	27	50	HYPOTHETICAL 24.6 KD PROTEIN IN CCPA 3'REGION (ORF2) [Bacillus megaterium]
1509	Bt1Gc2648	Bt1G3900	549-233	g267497	146	154	2.80E-10	37	23	HYPOTHETICAL 50.9 KD PROTEIN IN KATA 3'REGION (ORF A) [Bacillus firmus]
1510	Bt1Gc2652	Bt1G3901	1-180	g2116973	135	169	3.50E-12	62	17	(D87979) YfmK [Bacillus subtilis]
1511	Bt1Gc2656	Bt1G3902	424-1	g2619006	240	284	1.10E-24	40	30	(AF027868) putative L-amino acid oxidase precursor [Bacillus subtilis]
1512	Bt1Gc2677	Bt1G3903	1-215	g2635763	121	141	7.10E-09	50	15	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1513	Bt1Gc2681	Bt1G3904	185-1	g3183561	112	145	6.30E-10	56	23	VEGETATIVE PROTEIN 296 (VEG296) [Bacillus subtilis]
1514	Bt1Gc2678	Bt1G3905	1-255	g225559	214	237	2.70E-19	54	18	ORF IS231C [Bacillus thuringiensis]
1515	Bt1Gc2685	Bt1G3906	19-265	g4584149	76	77	0.032	39	26	(AJ010111) cytochrome aa3 controlling protein [Bacillus cereus]
1516	Bt1Gc2696	Bt1G3907	1-238	g2337799	97	125	3.10E-07	36	19	(Y13937) putative Dfp protein [Bacillus subtilis]
1517	Bt1Gc2679	Bt1G3908	7-248	g4584150	155	177	2.70E-13	48	26	(AJ010111) cytochrome caa3 oxidase assembly factor [Bacillus cereus]
1518	Bt1Gc2700	Bt1G3909	1-339	g2494761	159	224	1.00E-17	41	20	CTP SYNTHASE (UTP--AMMONIA LIGASE) (CTP SYNTHETASE) [Synecocystis sp.]
1519	Bt1Gc2725	Bt1G3911	220-440	g2634068	149	87	0.0055	55	14	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
1520	Bt1Gc2733	Bt1G3912	348-1	g1769944	88	143	5.50E-09	35	21	(X98626) SNF2 [Bacillus cereus]
1521	Bt1Gc2746	Bt1G3913	353-1	g1770026	206	226	4.20E-18	45	25	(Z75208) hypothetical protein [Bacillus subtilis]
1522	Bt1Gc2734	Bt1G3914	1-136	g732324	99	112	7.00E-06	48	12	PUTATIVE UDP-N-ACETYLGLUCOSAMINE 2- EPIMERASE (UDP-GLCNAC-2-EPIMERASE) [Bacillus subtilis]
1523	Bt1Gc2748	Bt1G3915	188-1	g1934606	121	141	4.90E-09	46	16	(U93874) cystathionine gamma-lyase [Bacillus subtilis]
1524	Bt1Gc2747	Bt1G3916	1-159	g585920	153	156	6.60E-10	62	4	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Bacillus subtilis]
1525	Bt1Gc2755	Bt1G3917	1-410	g3915673	118	182	1.20E-12	35	11	PROBABLE SERINE ACTIVATING ENZYME [Bacillus subtilis]
1526	Bt1Gc2765	Bt1G3918	397-1	g585375	126	176	2.70E-13	35	45	DIMETHYLADENOSINE TRANSFERASE (S- ADENOSYLMETHIONINE-6-N', N'-ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RRNA

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1527	Bt1Gc2770	Bt1G3919	212-108	g2635181	110	110	8.00E-06	60	11	DIMETHYLASE) (HIGH LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN DIMETHYLTRANSFERASE) [Bacillus subtilis]
1528	Bt1Gc2782	Bt1G3920	354-1	g2154723	343	271	1.50E-23	53	42	(Z99117) similar to protease [Bacillus subtilis]
1529	Bt1Gc2766	Bt1G3921	182-1	g2149596	198	216	1.00E-16	69	9	(Y11193) transcription antiterminator [Bacillus subtilis]
1530	Bt1Gc2789	Bt1G3922	371-1	g135175	270	246	6.00E-20	50	19	(U89796) chitinase [Bacillus thuringiensis]
1531	Bt1Gc2791	Bt1G3923	452-1	g140338	160	192	3.40E-15	37	70	THREONYL-TRNA SYNTHETASE 1 (THREONINE--TRNA LIGASE) (THRRS) [Bacillus subtilis]
1532	Bt1Gc2797	Bt1G3924	1-182	g2791497	172	199	2.90E-15	68	14	HYPOTHETICAL 23.1 KD PROTEIN IN DMSC-PFLA INTERGENIC REGION [Escherichia coli]
1533	Bt1Gc2824	Bt1G3925	145-1	g4584097	120	138	5.50E-09	61	17	(AL021246) clpX [Mycobacterium tuberculosis]
1534	Bt1Gc2840	Bt1G3926	1-315	g2982867	77	108	1.70E-05	28	29	(AJ010131) yfkH [Bacillus cereus]
1535	Bt1Gc2835	Bt1G3927	1-212	g121357	125	166	1.30E-11	49	16	(AE000675) riboflavin specific deaminase [Aquifex aeolicus]
1536	Bt1Gc2842	Bt1G3928	505-1	g732300	180	161	6.60E-12	30	68	GLUTAMINE SYNTHETASE (GLUTAMATE--AMMONIA LIGASE) [Bacillus cereus]
1537	Bt1Gc2843	Bt1G3929	1-351	g1093432	77	107	3.60E-05	26	24	HYPOTHETICAL 27.6 KD PROTEIN IN ACUC 5'REGION (ORFB) [Bacillus subtilis]
1538	Bt1Gc2850	Bt1G3930	433-1	g3183506	132	202	3.00E-16	35	52	catalase [Helianthus annuus]
1539	Bt1Gc2841	Bt1G3931	1-437	g1750108	525	416	6.30E-39	69	43	HYPOTHETICAL 31.7 KD PROTEIN IN GLTP-CWLJ INTERGENIC REGION (ORF13) [Bacillus subtilis]
1540	Bt1Gc2862	Bt1G3932	305-1	g1168772	139	158	3.50E-10	40	9	(U66480) YnbA [Bacillus subtilis]
1541	Bt1Gc2859	Bt1G3933	357-1	g1881355	120	141	1.20E-09	31	50	CARBAMOYL-PHOSPHATE SYNTHASE, PYRIMIDINE-SPECIFIC, LARGE CHAIN (CARBAMOYL-PHOSPHATE SYNTHETASE AMMONIA CHAIN) [Bacillus caldolyticus]
1542	Bt1Gc2865	Bt1G3934	145-1	g2117582	133	112	3.30E-05	63	4	(AB001488) FUNCTION UNKNOWN, SIMILAR PRODUCT IN SYNECHOCYSTIS. [Bacillus subtilis]
1543	Bt1Gc2878	Bt1G3935	357-18	g321919	79	143	5.30E-10	34	75	respiratory nitrate reductase alpha chain - Bacillus subtilis [Bacillus subtilis]
1544	Bt1Gc2913	Bt1G3937	1-372	g531829	120	161	6.60E-12	38	81	hypothetical 16.9K protein - Salmonella typhimurium plasmid NTP16 []
1545	Bt1Gc2910	Bt1G3938	770-58	g436923	125	49	0.28	40	91	(U12390) beta-galactosidase alpha peptide [cloning vector pSport1]
										(U01849) ORF1 [Trypanosoma brucei]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1546	Bt1Gc2918	Bt1G3939	274-1	g118672	128	141	7.30E-09	38	19	LIPOAMIDE DEHYDROGENASE COMPONENT OF PYRUVATE DEHYDROGENASE COMPLEX (E3) (DIHYDROLIPOAMIDE DEHYDROGENASE) (S COMPLEX, 50 KD SUBUNIT) [Bacillus subtilis]
1547	Bt1Gc2923	Bt1G3940	514-1	g2914486	863	889	4.70E-89	97	57	Phosphatidylinositol-Specific Phospholipase C Mutant D198e []
1548	Bt1Gc2925	Bt1G3941	24-414	g3256661	127	155	2.90E-11	29	93	(AP000001) 136aa long hypothetical protein [Pyrococcus horikoshii]
1549	Bt1Gc2927	Bt1G3942	400-103	g2497400	92	120	2.30E-07	28	46	HYPOTHETICAL TRANSPOSASE-LIKE PROTEIN HI1721 [Haemophilus influenzae Rd]
1550	Bt1Gc2928	Bt1G3943	395-1	g3122219	261	259	6.10E-22	41	31	HISTIDINOL DEHYDROGENASE (HDH) [Bacillus subtilis]
1551	Bt1Gc2997	Bt1G3944	1-173	g399870	128	150	3.50E-10	53	18	DELTA-AMINOLEVULINIC ACID DEHYDRATASE (PORPHOBILINOGEN SYNTHASE) (ALAD) (ALADH) [Bacillus subtilis]
1551	Bt1Gc2997	Bt1G3945	188-423	g399784	133	149	8.50E-10	43	18	GLUTAMATE-1-SEMIALDEHYDE 2,1-AMINOMUTASE (GSA) (GLUTAMATE-1-SEMIALDEHYDE AMINOTRANSFERASE) (GSA-AT) [Bacillus subtilis]
1552	Bt1Gc3013	Bt1G3948	288-1	g2507017	210	200	1.50E-15	51	25	HEMOLYSIN BL BINDING COMPONENT PRECURSOR (ENTEROTOXIN 40 KD SUBUNIT) [Bacillus cereus]
1553	Bt1Gc3012	Bt1G3950	97-400	g1945050	129	128	1.70E-07	38	23	(U63928) L2 protein [Bacillus cereus]
1554	Bt1Gc3034	Bt1G3951	1-505	g3023262	670	690	5.80E-68	73	43	PROBABLE NADH-DEPENDENT BUTANOL DEHYDROGENASE 1 [Bacillus subtilis]
1555	Bt1Gc3033	Bt1G3952	293-1	g267052	139	158	1.40E-11	45	41	UROPORPHYRIN-III C-METHYLTRANSFERASE (UROGEN III METHYLASE) (SUMT) (UROPORPHYRINOGEN III METHYLASE) (UROM) [Bacillus megaterium]
1556	Bt1Gc3043	Bt1G3953	329-1	g586866	155	177	9.60E-13	41	23	HYPOTHETICAL 53.2 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION [Bacillus subtilis]
1557	Bt1Gc3046	Bt1G3954	324-1	g586866	86	118	2.30E-06	40	22	HYPOTHETICAL 53.2 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION [Bacillus subtilis]
1558	Bt1Gc3049	Bt1G3955	331-1	g586866	157	170	5.50E-12	42	23	HYPOTHETICAL 53.2 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION [Bacillus subtilis]
1559	Bt1Gc3054	Bt1G3956	290-1	g1731087	119	169	3.20E-12	35	28	HYPOTHETICAL 39.0 KD PROTEIN IN GLNQ-ANSR INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1560	Bt1Gc3055	Bt1G3957	1-251	g1894742	117	133	1.00E-08	37	35	(Z92952) product similar to Staphylococcus aureus CapB protein [Bacillus subtilis]
1561	Bt1Gc3057	Bt1G3958	379-1	g473955	285	313	5.20E-28	52	30	(D29674) DNA-binding protein [Lactobacillus sp.]
1562	Bt1Gc3060	Bt1G3959	1-486	g1805397	447	353	3.00E-32	57	55	(D50453) homologue of unidentified protein of E. coli [Bacillus subtilis]
1563	Bt1Gc3070	Bt1G3960	191-1	g2293164	133	155	1.60E-10	48	16	(AF008220) SAM synthase [Bacillus subtilis]
1564	Bt1Gc3087	Bt1G3961	1-300	g1703066	112	156	1.10E-10	36	26	ACYL-COA DEHYDROGENASE, SHORT-CHAIN SPECIFIC (SCAD) (BUTYRYL-COA DEHYDROGENASE) [Clostridium acetobutylicum]
1565	Bt1Gc3088	Bt1G3962	1-398	g2226226	203	234	5.30E-19	39	29	(Y14083) Hypothetical protein [Bacillus subtilis]
1566	Bt1Gc3096	Bt1G3963	1-272	g68538	238	257	7.10E-21	56	10	valine--tRNA ligase (EC 6.1.1.9) - Bacillus stearothermophilus []
1567	Bt1Gc3100	Bt1G3964	1-379	g126054	187	241	2.20E-20	42	39	L-LACTATE DEHYDROGENASE X []
1568	Bt1Gc3101	Bt1G3965	396-8	g2982937	119	164	3.20E-12	26	77	(AE000679) superoxide dismutase (Cu/Zn) [Aquifex aeolicus]
1569	Bt1Gc3086	Bt1G3966	219-1	g732334	115	149	7.20E-10	43	18	HYPOTHETICAL 44.4 KD PROTEIN IN EPR-GALK INTERGENIC REGION [Bacillus subtilis]
1570	Bt1Gc3098	Bt1G3967	205-1	g3122129	145	156	8.60E-11	49	20	GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE 2 (GAPDH) [Bacillus subtilis]
1571	Bt1Gc3116	Bt1G3968	1-395	g4584142	238	288	2.30E-25	50	46	(AJ010139) DNA alkylation repair enzyme [Bacillus cereus]
1572	Bt1Gc3115	Bt1G3969	1-387	g2633128	312	323	4.50E-29	46	86	(Z99108) yfjM [Bacillus subtilis]
1573	Bt1Gc3132	Bt1G3970	268-1	g2497382	153	179	4.60E-13	43	21	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1574	Bt1Gc3134	Bt1G3971	410-1	g1350849	290	309	3.10E-26	47	12	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Staphylococcus aureus]
1575	Bt1Gc3135	Bt1G3972	328-114	g225559	86	119	1.80E-06	43	17	ORF IS231C [Bacillus thuringiensis]
1576	Bt1Gc3140	Bt1G3973	269-103	g2331287	81	101	0.0001	39	15	(AF013188) release factor 2 [Bacillus subtilis]
1577	Bt1Gc3146	Bt1G3974	1-248	g4584201	140	143	5.30E-10	41	60	(AJ000394) hypothetical protein [Bacillus cereus]
1578	Bt1Gc3152	Bt1G3975	334-1	g1731052	206	240	2.60E-19	53	14	PROBABLE 1-DEOXYXYLULOSE-5-PHOSPHATE SYNTHASE (DXP SYNTHASE) [Bacillus subtilis]
1579	Bt1Gc3149	Bt1G3976	364-1	g4204103	82	122	7.80E-07	28	26	(U10879) PtsW; cell division protein [Streptomyces coelicolor]
1580	Bt1Gc3156	Bt1G3977	1-271	g2116974	100	144	1.10E-08	39	8	(D87979) YfmJ [Bacillus subtilis]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
1581	Bt1Gc3183	Bt1G3978	1-126	g1708794	145	154	4.40E-10	76	7	GTP-BINDING PROTEIN LEPA [Bacillus subtilis]
1582	Bt1Gc3191	Bt1G3979	1-182	g3123196	81	117	1.90E-06	47	16	CITRATE SYNTHASE II [Bacillus subtilis]
1583	Bt1Gc3192	Bt1G3980	459-1	g2632017	343	383	2.00E-35	49	28	(AJ002571) DppE [Bacillus subtilis]
1584	Bt1Gc3198	Bt1G3981	14-266	g585920	140	173	1.00E-11	48	7	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Bacillus subtilis]
1585	Bt1Gc3208	Bt1G3982	391-54	g2636153	74	108	6.40E-05	30	12	(Z99122) similar to SNF2 helicase [Bacillus subtilis]
1586	Bt1Gc3210	Bt1G3983	1-233	g1881335	294	310	1.10E-27	74	27	(AB001488) SIMILAR TO YQFU, YXKD, YITB OF B. SUBTILIS. [Bacillus subtilis]
1587	Bt1Gc3211	Bt1G3984	354-1	g2108273	148	173	1.50E-12	34	31	(X92868) NAD alcohol dehydrogenase [Bacillus subtilis]
1588	Bt1Gc3214	Bt1G3985	282-39	g113739	134	124	1.30E-07	46	33	METHIONINE AMINOPEPTIDASE (MAP) (PEPTIDASE M) [Bacillus subtilis]
1589	Bt1Gc3207	Bt1G3986	295-1	g2108273	139	198	2.60E-15	42	27	(X92868) NAD alcohol dehydrogenase [Bacillus subtilis]
1590	Bt1Gc3227	Bt1G3987	293-39	g4584135	258	111	1.80E-06	68	43	(AJ010138) stage II sporulation protein [Bacillus cereus]
1591	Bt1Gc3234	Bt1G3988	1-329	g4584201	179	190	5.60E-15	39	80	(AJ000394) hypothetical protein [Bacillus cereus]
1592	Bt1Gc3248	Bt1G3989	368-1	g1710069	104	105	3.90E-05	27	33	RECF PROTEIN I
1593	Bt1Gc3251	Bt1G3990	107-250	g1770039	119	125	4.30E-08	56	24	(Z75208) hypothetical protein [Bacillus subtilis]
1594	Bt1Gc3252	Bt1G3991	362-1	g1709569	195	213	2.00E-17	40	42	PANTOATE--BETA-ALANINE LIGASE (PANTOTHENATE SYNTHETASE) (PANTOATE ACTIVATING ENZYME) [Bacillus subtilis]
1595	Bt1Gc3258	Bt1G3992	1-328	g2443256	96	131	1.00E-07	26	21	(D86417) YfmM [Bacillus subtilis]
1596	Bt1Gc3295	Bt1G3993	341-1	g2633124	247	265	6.30E-23	45	36	(Z99108) similar to divalent cation transport protein [Bacillus subtilis]
1597	Bt1Gc3298	Bt1G3994	433-235	g2635191	88	128	2.70E-08	46	31	(Z99117) glutamine ABC transporter (integral membrane protein) [Bacillus subtilis]
1598	Bt1Gc3300	Bt1G3995	478-300	g2501426	144	124	4.40E-07	55	14	PYRIMIDINE-NUCLEOSIDE PHOSPHORYLASE [Bacillus stearothermophilus]
1599	Bt1Gc3305	Bt1G3996	1-230	g629038	115	136	2.90E-08	40	15	hypothetical protein - Bacillus subtilis I
1600	Bt1Gc3310	Bt1G3997	469-95	g1169919	324	263	7.50E-22	56	21	GLUCOSAMINE--FRUCTOSE-6-PHOSPHATE AMINOTRANSFERASE (ISOMERIZING) (HEXOSEPHOSPHATE AMINOTRANSFERASE) (FRUCTOSE-6-PHOSPHATE AMIDOTRANSFERASE) (GFAT) (L-GLUTAMINE-D-FRUCTOSE-6-PHOSPHATE AMIDOTRANSFERASE) (GLUCOSAMINE-6-PH.. [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1601	Bt1Gc3313	Bt1G3998	1-184	g629265	99	103	2.70E-05	45	26	aldehyde dehydrogenase - Mycoplasma capricolum (SGC3) (fragment) [Mycoplasma capricolum]
1602	Bt1Gc3326	Bt1G3999	227-1	g2415403	152	181	2.90E-13	49	17	(AF015775) acetylornithine deacetylase [Bacillus subtilis]
1603	Bt1Gc3333	Bt1G4000	380-1	g2462099	285	303	5.90E-27	52	55	(Y10908) sulfate transport ATP-binding protein [Bacillus cereus]
1604	Bt1Gc3335	Bt1G4001	1-244	g2226174	85	127	2.50E-07	33	17	(Y14081) hypothetical protein [Bacillus subtilis]
1605	Bt1Gc3353	Bt1G4002	359-1	g2274944	134	153	2.70E-10	33	30	(AJ000346) NapC protein [Enterococcus hirae]
1606	Bt1Gc3359	Bt1G4003	501-7	g3915568	185	204	5.20E-16	33	44	HYPOTHETICAL 39.4 KD OXIDOREDUCTASE IN HOM-MRGA INTERGENIC REGION [Bacillus subtilis]
1607	Bt1Gc3360	Bt1G4004	290-1	g148683	327	267	3.90E-23	63	52	(M55343) open reading frame [Frankia sp.]
1608	Bt1Gc3373	Bt1G4005	416-147	g118262	398	398	5.10E-37	90	99	DNA-BINDING PROTEIN II (HB) (HU) [Bacillus stearothermophilus]
1609	Bt1Gc3378	Bt1G4006	1-294	g3915989	189	247	8.70E-21	58	24	HYPOTHETICAL 45.3 KD PROTEIN IN PRKA-CSPB INTERGENIC REGION (ORF4) [Bacillus subtilis]
1610	Bt1Gc3382	Bt1G4007	127-1	g4063787	136	69	0.067	63	27	(AB008550) orf14; similar to S gene of P2:tail completion [Pseudomonas aeruginosa phage phi CTX]
1611	Bt1Gc3396	Bt1G4009	1-223	g2497382	129	139	1.00E-08	49	17	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1612	Bt1Gc3412	Bt1G4010	1-379	g134396	263	132	1.60E-07	50	15	PREPROTEIN TRANSLOCASE SECA SUBUNIT [Bacillus subtilis]
1613	Bt1Gc3427	Bt1G4011	346-3	g129521	130	145	3.30E-10	33	59	PARA-AMINOBENZOATE SYNTHASE GLUTAMINE AMIDOTRANSFERASE COMPONENT II (ADC SYNTHASE) / ANTHRANILATE SYNTHASE COMPONENT II [Bacillus subtilis]
1614	Bt1Gc3430	Bt1G4012	1-339	g2293322	118	114	5.50E-06	33	25	(AF008220) branch-chain amino acid transporter [Bacillus subtilis]
1615	Bt1Gc3434	Bt1G4013	368-1	g2833392	188	177	1.30E-12	43	22	HYPOTHETICAL 61.5 KD PROTEIN IN ADEC-PDHA INTERGENIC REGION [Bacillus subtilis]
1616	Bt1Gc3445	Bt1G4014	442-1	g586866	177	203	1.40E-15	38	31	HYPOTHETICAL 53.2 KD PROTEIN IN XPAC-ABRB INTERGENIC REGION [Bacillus subtilis]
1617	Bt1Gc3450	Bt1G4015	321-42	g121190	139	184	2.40E-14	43	38	GLUCOSE INHIBITED DIVISION PROTEIN B [Bacillus subtilis]
1618	Bt1Gc3463	Bt1G4016	129-458	g1706797	249	189	1.70E-14	40	32	FERRICHRONE TRANSPORT PERMEASE PROTEIN FHUG [Bacillus subtilis]
1619	Bt1Gc3494	Bt1G4017	291-1	g2293328	134	175	2.70E-13	42	35	(AF008220) putative morphine dehydrogenase [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aaf_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1620	Bt1Gc3517	Bt1G4018	1-484	g2619056	450	432	1.30E-40	56	36	(AF027868) putative transporter [Bacillus subtilis]
1621	Bt1Gc3514	Bt1G4019	1-345	g2633471	256	270	2.60E-23	54	27	(Z99109) similar to 3-oxoacyl-acyl-carrier protein synthase [Bacillus subtilis]
1622	Bt1Gc3523	Bt1G4020	1-469	g466137	74	109	4.90E-06	34	38	HYPOTHETICAL 25.6 KD PROTEIN IN PLCB-LDH INTERGENIC REGION (ORFA) [Listeria monocytogenes]
1623	Bt1Gc3532	Bt1G4021	219-1	g994736	102	106	4.50E-06	40	77	(M18327) LacOPZ-alpha peptide from pUC9; putative [cloning vectors]
1624	Bt1Gc3551	Bt1G4022	372-1	g1169253	136	178	6.20E-13	38	28	ANAEROBIC C4-DICARBOXYLATE TRANSPORTER DCUB [Haemophilus influenzae Rd]
1625	Bt1Gc3553	Bt1G4023	1-411	g113346	132	190	1.50E-13	35	11	ATP-DEPENDENT NUCLEASE SUBUNIT B [Bacillus subtilis]
1626	Bt1Gc3555	Bt1G4024	337-1	g2127795	123	180	2.20E-13	38	31	carbamoyl-phosphate synthase, small chain (EC 6.3.-.-) - Methanococcus jannaschii []
1627	Bt1Gc3566	Bt1G4025	1-333	g1708794	312	349	3.30E-31	62	18	GTP-BINDING PROTEIN LEPA [Bacillus subtilis]
1628	Bt1Gc3574	Bt1G4026	1-210	g586885	180	137	2.10E-08	53	14	HYPOTHETICAL 56.1 KD PROTEIN IN MFD-DIVIC INTERGENIC REGION [Bacillus subtilis]
1629	Bt1Gc3603	Bt1G4027	245-1	g130904	148	79	0.00063	43	13	DNA PRIMASE [Bacillus subtilis]
1630	Bt1Gc3597	Bt1G4028	256-1	g1346795	98	153	5.80E-10	36	14	DNA PRIMASE [Listeria monocytogenes]
1631	Bt1Gc3652	Bt1G4029	212-464	g132246	137	154	1.80E-10	35	23	RECF PROTEIN [Bacillus subtilis]
1632	Bt1Gc3686	Bt1G4031	66-405	g3688823	187	196	1.30E-15	34	95	(AF084104) hypothetical protein [Bacillus firmus]
1633	Bt1Gc3690	Bt1G4032	258-1	g2632519	190	185	4.00E-14	51	26	(Z99105) similar to hypothetical proteins [Bacillus subtilis]
1634	Bt1Gc3709	Bt1G4034	1-301	g232186	107	127	3.40E-06	31	2	GRAMICIDIN S SYNTHETASE II (GRAMICIDIN S BIOSYNTHESIS GRSB PROTEIN) [Brevibacillus brevis]
1635	Bt1Gc3713	Bt1G4035	380-145	g2829479	276	281	1.30E-24	73	59	50S RIBOSOMAL PROTEIN L11 [Bacillus stearothermophilus, Peptide, 133 aa]
1636	Bt1Gc3721	Bt1G4036	1-310	g1934609	93	159	2.50E-10	38	10	(U93874) formate dehydrogenase chain A [Bacillus subtilis]
1637	Bt1Gc3719	Bt1G4037	1-278	g1169919	162	202	2.90E-15	47	15	GLUCOSAMINE--FRUCTOSE-6-PHOSPHATE AMINOTRANSFERASE (ISOMERIZING) (HEXOSEPHOSPHATE AMINOTRANSFERASE) (FRUCTOSE-6-PHOSPHATE AMIDOTRANSFERASE) (GFAT) (L-GLUTAMINE-D-FRUCTOSE-6-PHOSPHATE AMIDOTRANSFERASE) (GLUCOSAMINE-6-PH.. [Bacillus subtilis])
1638	Bt1Gc3743	Bt1G4038	530-89	g1731017	208	250	2.80E-20	43	20	HYPOTHETICAL 80.1 KD PROTEIN IN SODA-COMGA

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1639	Bt1Gc3747	Bt1G4039	1-135	g2290175	115	135	5.40E-08	61	7	INTERGENIC REGION [Bacillus subtilis] (AF067645) spore germination protein GerIA [Bacillus cereus]
1640	Bt1Gc3759	Bt1G4040	141-1	g1708641	103	129	7.10E-08	48	15	(U51115) YeaC [Bacillus subtilis]
1641	Bt1Gc3761	Bt1G4041	1-427	g4558012	178	173	1.60E-12	36	37	Chain A, Alanine Racemase With Bound Propionate Inhibitor []
1642	Bt1Gc3760	Bt1G4042	500-177	g2501575	121	140	4.00E-09	26	35	HYPOTHETICAL 33.3 KD PROTEIN SLL1263 [Synecocystis sp.]
1643	Bt1Gc3772	Bt1G4043	411-1	g1078767	238	266	7.60E-22	37	16	acetaldehyde dehydrogenase (acetylating) (EC 1.2.1.10) / alcohol dehydrogenase (EC 1.1.1.1) - Entamoeba histolytica [Entamoeba histolytica]
1644	Bt1Gc3779	Bt1G4044	1-556	g127481	203	259	2.70E-22	32	97	GTP CYCLOHYDROLASE I (GTP-CH-I) [Bacillus subtilis]
1645	Bt1Gc3785	Bt1G4045	1-272	g629038	141	126	3.50E-07	48	17	hypothetical protein - Bacillus subtilis []
1646	Bt1Gc3796	Bt1G4046	1-275	g732376	173	206	1.10E-16	47	36	HYPOTHETICAL 29.5 KD PROTEIN IN ROCC-PTA INTERGENIC REGION [Bacillus subtilis]
1647	Bt1Gc3798	Bt1G4047	262-1	g2635763	153	213	1.10E-16	44	18	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1648	Bt1Gc3817	Bt1G4049	290-1	g4584150	269	285	4.80E-25	59	31	(AJ010111) cytochrome caa3 oxidase assembly factor [Bacillus cereus]
1649	Bt1Gc3812	Bt1G4050	319-1	g994736	78	91	0.00017	54	57	(M18327) LacOPZ-alpha peptide from pUC9; putative [cloning vectors]
1650	Bt1Gc3819	Bt1G4051	267-1	g3345481	152	190	5.40E-14	44	15	(AB016285) phosphotransferase system enzyme I [Bacillus sp.]
1651	Bt1Gc3821	Bt1G4052	59-294	g1730925	178	186	1.50E-14	43	45	HYPOTHETICAL 19.0 KD PROTEIN IN ILVD-THYB INTERGENIC REGION [Bacillus subtilis]
1652	Bt1Gc3822	Bt1G4053	445-1	g1172699	376	378	1.30E-34	54	26	PHOSPHENOLPYRUVATE-PROTEIN PHOSPHOTRANSFERASE (PHOSPHOTRANSFERASE SYSTEM, ENZYME I) [Bacillus stearothermophilus]
1653	Bt1Gc3826	Bt1G4054	1-259	g4584148	210	225	6.70E-18	54	9	(AJ010111) pyruvate carboxylase [Bacillus cereus]
1654	Bt1Gc3830	Bt1G4055	270-1	g3025180	100	96	0.00048	40	20	HYPOTHETICAL 49.9 KD PROTEIN IN CITA-SSPB INTERGENIC REGION [Bacillus subtilis]
1655	Bt1Gc3841	Bt1G4057	1-503	g1945051	157	229	9.00E-19	35	43	(U63928) L1 protein [Bacillus cereus]
1656	Bt1Gc3844	Bt1G4058	518-1	g2293156	183	237	5.80E-20	36	91	(AF008220) YtiB [Bacillus subtilis]
1657	Bt1Gc3850	Bt1G4059	1-344	g321919	184	79	0.004	49	77	hypothetical 16.9K protein - Salmonella typhimurium plasmid NTP16 []

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1658	Bt1Gc3852	Bt1G4060	1-266	g3041750	259	269	1.60E-22	64	15	THIAMIN BIOSYNTHESIS PROTEIN THIC [Bacillus subtilis]
1659	Bt1Gc3855	Bt1G4061	1-341	g728801	174	157	9.20E-11	41	29	ACETON UTILIZATION ACUC PROTEIN [Bacillus subtilis]
1660	Bt1Gc3866	Bt1G4063	1-317	g1945051	152	166	9.30E-12	33	28	(U63928) L1 protein [Bacillus cereus]
1661	Bt1Gc3877	Bt1G4064	1-491	g126295	240	333	9.50E-29	42	13	LINE-1 REVERSE TRANSCRIPTASE HOMOLOG []
1662	Bt1Gc3870	Bt1G4065	79-490	g2648849	163	207	8.80E-17	40	56	(AE000986) 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid hydrolase (pcbD) [Archaeoglobus fulgidus]
1663	Bt1Gc3900	Bt1G4066	1-140	g3023940	101	125	4.30E-08	47	32	HIT PROTEIN [Bacillus subtilis]
1664	Bt1Gc3929	Bt1G4067	310-1	g4584097	224	208	6.90E-17	53	36	(AJ010131) yfkH [Bacillus cereus]
1665	Bt1Gc3931	Bt1G4068	1-407	g1881268	263	273	3.80E-23	48	26	(AB001488) ATP-DEPENDENT RNA HELICASE DEAD HOMOLOG. [Bacillus subtilis]
1666	Bt1Gc3917	Bt1G4069	1-206	g2149596	162	188	1.00E-13	51	11	(U89796) chitinase [Bacillus thuringiensis]
1667	Bt1Gc3953	Bt1G4070	330-1	g2635777	374	388	5.80E-36	66	92	(Z99120) similar to arsenate reductase [Bacillus subtilis]
1668	Bt1Gc3959	Bt1G4071	445-1	g1064900	161	179	8.20E-14	33	61	(X92970) orfB [Escherichia coli]
1669	Bt1Gc3960	Bt1G4072	301-1	g2117766	136	180	3.00E-13	42	25	orthine aminotransferase rocD - Bacillus subtilis []
1670	Bt1Gc3964	Bt1G4073	1-263	g1730252	153	171	2.50E-12	45	23	GCPE PROTEIN HOMOLOG [Bacillus subtilis]
1671	Bt1Gc3966	Bt1G4074	435-1	g2833392	155	204	1.50E-15	34	26	HYPOTHETICAL 61.5 KD PROTEIN IN ADEC-PDHA INTERGENIC REGION [Bacillus subtilis]
1672	Bt1Gc3969	Bt1G4075	238-1	g125524	402	400	3.10E-37	98	22	TRANSCRIPTIONAL REPRESSOR PROTEIN KORB [Enterobacteriaceae]
1673	Bt1Gc3970	Bt1G4076	403-252	g141263	201	219	4.70E-18	82	79	HYPOTHETICAL 6.8 KD PROTEIN IN REPLICATION ORIGIN REGION []
1674	Bt1Gc3974	Bt1G4077	251-1	g2633542	99	130	8.10E-08	36	21	(Z99110) similar to cystathionine beta-lyase [Bacillus subtilis]
1675	Bt1Gc3978	Bt1G4078	1-385	g133719	440	337	1.50E-30	73	97	30S RIBOSOMAL PROTEIN S11 (BS11) [Bacillus subtilis]
1676	Bt1Gc3981	Bt1G4079	1-559	g2636487	651	653	4.80E-64	67	47	(Z99124) pyrimidine-nucleoside transport protein [Bacillus subtilis]
1677	Bt1Gc3982	Bt1G4080	1-306	g3123214	117	154	9.20E-11	39	35	ELONGATION FACTOR TS (EF-TS) [Bacillus subtilis]
1678	Bt1Gc3983	Bt1G4081	282-88	g3183468	113	117	9.20E-07	42	25	HYPOTHETICAL 29.8 KD PROTEIN IN FRR-CDSA INTERGENIC REGION [Bacillus subtilis]
1679	Bt1Gc3998	Bt1G4082	1-187	g2982196	165	167	2.80E-10	52	1	(AF007865) bacitracin synthetase 3; BacC [Bacillus licheniformis]
1680	Bt1Gc4008	Bt1G4083	1-357	g2829689	168	185	9.80E-14	40	28	TRIGGER FACTOR (TF) (VEGETATIVE PROTEIN 2)

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1681	Bt1Gc4006	Bt1G4084	227-1	g2117582	148	174	8.20E-12	53	6	(VEG2) [Bacillus subtilis] respiratory nitrate reductase alpha chain - [Bacillus subtilis]
1682	Bt1Gc4011	Bt1G4085	316-1	g2633967	94	120	7.30E-07	32	32	[Bacillus subtilis] (Z99112) signal recognition particle (docking protein)
1683	Bt1Gc4023	Bt1G4086	464-1	g730399	159	210	4.40E-16	36	25	PRKA PROTEIN [Bacillus subtilis]
1684	Bt1Gc4013	Bt1G4087	1-501	g2633966	170	177	7.70E-13	32	14	(Z99112) chromosome segregation SMC protein homolog
1685	Bt1Gc4027	Bt1G4088	1-254	g4512347	201	204	1.10E-15	58	18	[Bacillus subtilis] (AB011836) similar to B.subtilis yerN gene(87% identity)
1686	Bt1Gc4033	Bt1G4089	108-476	g3913859	162	63	7.80E-05	39	48	[Bacillus halodurans] HISF PROTEIN (CYCLASE) [Bacillus subtilis]
1687	Bt1Gc4039	Bt1G4090	1-284	g4138732	95	49	0.42	27	25	(Y17332) proline-rich protein [Zea mays]
1688	Bt1Gc4040	Bt1G4091	1-419	g2634076	223	229	6.80E-18	37	16	(Z99112) mutS [Bacillus subtilis]
1689	Bt1Gc4055	Bt1G4093	446-1	g4218500	230	281	1.60E-23	43	19	(Y16739) formate C-acetyltransferase [Piromyces sp.]
1690	Bt1Gc4060	Bt1G4094	1-203	g1709698	177	194	2.80E-14	63	10	POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE (POLYNUCLEOTIDE PHOSPHORYLASE) (PNPASE) (VEGETATIVE PROTEIN 15) (VEG15) [Bacillus subtilis]
1691	Bt1Gc4075	Bt1G4095	348-1	g543864	160	204	3.40E-15	33	12	MG(2+) TRANSPORT ATPASE, P-TYPE 1 [Salmonella typhimurium]
1692	Bt1Gc4077	Bt1G4096	1-175	g1173058	237	234	1.20E-19	73	32	50S RIBOSOMAL PROTEIN L6 (BL10) [Bacillus subtilis]
1693	Bt1Gc4090	Bt1G4097	360-108	g1174408	195	134	4.80E-09	56	98	STAGE V SPORULATION PROTEIN S [Bacillus subtilis]
1694	Bt1Gc4091	Bt1G4098	407-1	g732340	116	145	7.40E-10	31	50	HYDROXYETHYLTHIAZOLE KINASE (4-METHYL-5- BETA-HYDROXYETHYLTHIAZOLE KINASE) (THZ KINASE) (TH KINASE) [Bacillus subtilis]
1695	Bt1Gc4097	Bt1G4099	299-1	g1177035	112	118	9.60E-07	34	34	PROBABLE ABC TRANSPORTER PERMEASE PROTEIN IN SODA-COMGA INTERGENIC REGION (ORF73) [Bacillus subtilis]
1696	Bt1Gc4132	Bt1G4100	1-317	g2226165	172	170	9.80E-12	42	16	(Y14080) hypothetical protein [Bacillus subtilis]
1697	Bt1Gc4136	Bt1G4101	501-187	g3183454	208	207	8.80E-17	48	36	HYPOTHETICAL 30.7 KD PROTEIN IN MCPC-KINA INTERGENIC REGION [Bacillus subtilis]
1698	Bt1Gc4135	Bt1G4102	304-1	g2497382	177	220	1.50E-17	44	24	TRANSPORPOSE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1699	Bt1Gc4146	Bt1G4103	1-410	g728788	303	353	8.30E-32	51	24	ACETYL-COENZYME A SYNTHETASE (ACETATE-- COA LIGASE) (ACYL-ACTIVATING ENZYME)

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1700	Bt1Gc4142	Bt1G4104	123-395	g144312	106	114	6.30E-07	39	74	(ACETYL-COA SYNTHASE) [Bacillus subtilis]
1701	Bt1Gc4150	Bt1G4105	1-248	g121516	105	136	2.90E-08	35	16	(J01566) 13.8 kd ORF [Plasmid ColE1] GLUCONOKINASE (GLUCONATE KINASE) [Bacillus subtilis]
1702	Bt1Gc4144	Bt1G4106	1-299	g1168772	146	201	9.10E-15	46	9	CARBAMOYL-PHOSPHATE SYNTHASE, PYRIMIDINE-SPECIFIC, LARGE CHAIN (CARBAMOYL-PHOSPHATE SYNTHETASE AMMONIA CHAIN) [Bacillus caldolyticus]
1703	Bt1Gc4153	Bt1G4107	293-1	g2072373	107	137	1.00E-08	35	29	(Y09252) phosphoribosylaminoimidazolecarboxamide formyltransferase [Bacillus cereus]
1704	Bt1Gc4154	Bt1G4108	242-459	g123765	94	141	9.50E-09	51	13	UROCANATE HYDRATASE (UROCANASE) (IMIDAZOLONEPROPIONATE HYDROLASE) [Pseudomonas putida]
1705	Bt1Gc4160	Bt1G4109	189-1	g123287	98	150	2.10E-09	48	7	ISOLEUCYL-TRNA SYNTHETASE (ISOLEUCINE-- TRNA LIGASE) (ILERS) [Bacillus subtilis]
1706	Bt1Gc4172	Bt1G4110	539-41	g3025181	252	203	1.30E-15	42	35	HYPOTHETICAL 51.5 KD PROTEIN IN CITA-SSPB INTERGENIC REGION [Bacillus subtilis]
1707	Bt1Gc4174	Bt1G4111	313-1	g3914289	163	194	2.50E-14	41	16	TOPOISOMERASE IV SUBUNIT B [Bacillus subtilis]
1708	Bt1Gc4177	Bt1G4112	1-204	g585920	139	135	1.10E-07	46	6	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Bacillus subtilis]
1709	Bt1Gc4179	Bt1G4113	378-1	g584808	218	192	2.20E-14	43	27	ATP SYNTHASE BETA CHAIN (VEGETATIVE PROTEIN 31) (VEG31) [Bacillus subtilis]
1710	Bt1Gc4188	Bt1G4114	197-1	g2635763	132	150	7.60E-10	69	14	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1711	Bt1Gc4194	Bt1G4115	1-213	g3123297	112	141	1.20E-08	40	11	EXCINUCLEASE ABC SUBUNIT B (DNA PROTEIN) [Bacillus subtilis]
1712	Bt1Gc4198	Bt1G4116	207-1	g141329	97	112	3.20E-06	36	27	DIHYDROOROTATE DEHYDROGENASE ELECTRON TRANSFER SUBUNIT [Bacillus subtilis]
1713	Bt1Gc4200	Bt1G4117	1-254	g1853968	131	170	7.90E-13	45	30	(D88121) CPRD12 protein [Vigna unguiculata]
1714	Bt1Gc4193	Bt1G4118	164-1	g2492564	114	132	1.80E-07	52	6	EXCINUCLEASE ABC SUBUNIT A [Thermus thermophilus]
1715	Bt1Gc4201	Bt1G4119	1-235	g137018	367	381	3.20E-35	90	60	HYPOTHETICAL PROTEIN UL127 [human herpesvirus 5]
1716	Bt1Gc4203	Bt1G4120	380-1	g1731040	142	77	0.069	35	23	HYPOTHETICAL HELICASE IN SINI-GCVT INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1717	Bt1Gc4232	Bt1G4121	249-1	g1805393	127	147	1.90E-09	42	16	(D50453) 68% identity protein to 1-pyrroline-5-carboxylate dehydrogenase of B. subtilis [Bacillus subtilis]
1718	Bt1Gc4228	Bt1G4122	183-1	g3122350	113	130	3.50E-08	42	23	PROLIPOPROTEIN DIACYLGLYCERYL TRANSFERASE (SPORE GERMINATION PROTEIN GERF) [Bacillus subtilis]
1719	Bt1Gc4238	Bt1G4123	1-288	g2462959	191	177	1.30E-13	38	48	(AJ000974) putative adenosine 5-phosphosulfate kinase [Bacillus subtilis]
1720	Bt1Gc4241	Bt1G4124	231-29	g2497856	116	131	4.00E-08	41	21	MALATE DEHYDROGENASE [Bacillus israeli]
1721	Bt1Gc4240	Bt1G4125	1-253	g133466	99	114	4.10E-06	33	23	RNA POLYMERASE SIGMA FACTOR RPOD (SIGMA-A) (SIGMA-43) [Bacillus subtilis]
1722	Bt1Gc4254	Bt1G4127	280-1	g2497392	181	199	6.20E-16	49	37	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
1723	Bt1Gc4255	Bt1G4128	79-248	g3915204	123	160	1.90E-10	61	6	EXCINUCLEASE ABC SUBUNIT A [Bacillus subtilis]
1724	Bt1Gc4253	Bt1G4129	1-235	g595768	111	118	2.40E-07	42	47	(U13866) non-functional lacZ alpha peptide [Cloning vector]
1725	Bt1Gc4264	Bt1G4130	187-1	g2497382	115	159	6.90E-11	52	14	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1726	Bt1Gc4262	Bt1G4131	352-1	g2635778	154	180	6.70E-13	39	20	(Z99120) similar to butyryl-CoA dehydrogenase [Bacillus subtilis]
1727	Bt1Gc4263	Bt1G4132	242-1	g417332	87	127	2.60E-08	34	51	MUTT-LIKE PROTEIN (ORF154) [Streptomyces ambofaciens]
1728	Bt1Gc4301	Bt1G4133	1-314	g3290177	133	175	7.90E-13	38	30	(AF067645) spore germination protein GerIC [Bacillus cereus]
1729	Bt1Gc4281	Bt1G4134	355-1	g135811	153	200	1.20E-15	44	34	THREONINE SYNTHASE [Bacillus sp.]
1730	Bt1Gc4312	Bt1G4135	107-197	g3183561	90	108	9.30E-06	68	11	VEGETATIVE PROTEIN 296 (VEG296) [Bacillus subtilis]
1731	Bt1Gc4332	Bt1G4136	1-229	g2239276	386	399	4.00E-37	95	13	(L77763) neutral protease A [Bacillus thuringiensis]
1732	Bt1Gc4337	Bt1G4137	237-1	g1172465	320	322	5.70E-29	71	40	CDP-DIACYLGLYCEROL--GLYCEROL-3-PHOSPHATE 3-PHOSPHATIDYLTRANSFERASE (PHOSPHATIDYLGLYCEROPHOSPHATE SYNTHASE) (PGP SYNTHASE) [Bacillus subtilis]
1733	Bt1Gc4340	Bt1G4138	190-319	g2633489	75	115	1.60E-06	41	18	(Z99110) yjaZ [Bacillus subtilis]
1734	Bt1Gc4341	Bt1G4139	1-287	g585481	252	212	6.90E-16	57	8	TRANSCRIPTION-REPAIR COUPLING FACTOR (TRCF) [Bacillus subtilis]
1735	Bt1Gc4351	Bt1G4140	1-219	g2500053	121	148	1.20E-09	45	16	DNA REPAIR PROTEIN RADA HOMOLOG (DNA REPAIR PROTEIN SMS HOMOLOG) [Listeria]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1736	Bt1Gc4355	Bt1G4141	1-278	g1168646	133	132	2.90E-08	41	31	monocytogenes] GTP-BINDING PROTEIN ERA HOMOLOG (BEX PROTEIN) [Bacillus subtilis]
1737	Bt1Gc4354	Bt1G4143	309-1	g2829689	221	241	6.40E-20	47	24	TRIGGER FACTOR (TF) (VEGETATIVE PROTEIN 2) (VEG2) [Bacillus subtilis]
1738	Bt1Gc4358	Bt1G4144	1-264	g3329623	108	66	0.43	27	26	(AF078790) No definition line found [Caenorhabditis elegans]
1739	Bt1Gc4360	Bt1G4145	379-1	g2497382	270	276	7.20E-24	49	29	TRANSPPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1740	Bt1Gc4368	Bt1G4146	1-403	g267497	137	186	9.30E-14	33	26	HYPOTHETICAL 50.9 KD PROTEIN IN KATA 3 REGION (ORF A) [Bacillus firmus]
1741	Bt1Gc4367	Bt1G4147	1-311	g3777577	184	110	6.10E-06	47	38	(AF083408) kanamycin 3'-phosphotransferase [Cloning vector p34S-Km3]
1742	Bt1Gc4370	Bt1G4148	320-1	g1881268	129	117	3.20E-06	36	21	(AB001488) ATP-DEPENDENT RNA HELICASE DEAD HOMOLOG. [Bacillus subtilis]
1743	Bt1Gc4381	Bt1G4149	1-333	g1168649	182	213	2.00E-17	43	44	MENAQUINOL-CYTOCHROME C REDUCTASE CYTOCHROME B/C SUBUNIT [Bacillus subtilis]
1744	Bt1Gc4382	Bt1G4150	345-32	g1934806	104	140	1.20E-09	39	47	(Z93936) unknown [Bacillus subtilis]
1745	Bt1Gc4384	Bt1G4151	1-389	g2634887	198	217	7.70E-18	40	46	(Z99116) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1746	Bt1Gc4391	Bt1G4152	388-28	g584771	91	112	8.00E-06	30	29	GLUTAMATE N-ACETYLTRANSFERASE (ORNITHINE ACETYLTRANSFERASE) (ORNITHINE TRANSACETYLASE) (OATASE) / AMINO-ACID ACETYLTRANSFERASE (N-ACETYLGLUTAMATE SYNTHASE) (AGS) [Bacillus stearothermophilus]
1747	Bt1Gc4407	Bt1G4153	132-1	g1711644	115	134	4.80E-08	62	9	PROBABLE ASPARAGINYL-TRNA SYNTHETASE (ASPARAGINE--TRNA LIGASE) (ASNRS) [Synecocystis sp.]
1748	Bt1Gc4416	Bt1G4154	1-102	g3892204	154	105	0.00015	94	3	(AF078105) site-1 protease of sterol regulatory element binding proteins [Cricetulus griseus]
1749	Bt1Gc4432	Bt1G4155	286-1	g2635778	67	116	5.10E-06	37	17	(Z99120) similar to butyryl-CoA dehydrogenase [Bacillus subtilis]
1750	Bt1Gc4436	Bt1G4156	1-367	g2293213	479	483	5.00E-46	75	61	(AF008220) YtpR [Bacillus subtilis]
1751	Bt1Gc4441	Bt1G4157	1-302	g2632593	175	112	8.50E-06	41	23	(Z99105) multidrug-efflux transporter [Bacillus subtilis]
1752	Bt1Gc4443	Bt1G4158	474-217	g730103	145	152	2.40E-10	36	25	NARA PROTEIN [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1753	Bt1Gc4447	Bt1G4159	290-1	g1708474	158	180	4.80E-13	48	19	INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE (IMP DEHYDROGENASE) (IMPDH) (IMPD) [Streptococcus pyogenes]
1754	Bt1Gc4445	Bt1G4160	339-1	g2293149	110	146	2.20E-09	35	24	OSB-CoA synthase [Bacillus subtilis]
1755	Bt1Gc4472	Bt1G4161	1-263	g2851477	106	117	3.60E-06	35	16	ARGINYL-TRNA SYNTHETASE (ARGININE--TRNA LIGASE) (ARGRS) [Bacillus subtilis]
1756	Bt1Gc4485	Bt1G4162	1-332	g3183527	93	159	1.10E-11	37	67	TRANSCRIPTION ELONGATION FACTOR GRE (TRANSCRIPT CLEAVAGE FACTOR GRE) (GENERAL STRESS PROTEIN 20M) (GSP20M) [Bacillus subtilis]
1757	Bt1Gc4497	Bt1G4164	1-348	g1769947	241	265	1.30E-21	52	11	SNF [Bacillus cereus]
1758	Bt1Gc4504	Bt1G4165	1-398	g1881262	249	262	1.30E-22	41	35	(AB001488) SIMILAR TO ACYL-COA DEHYDROGENASE. [Bacillus subtilis]
1759	Bt1Gc4522	Bt1G4166	238-1	g3687417	149	172	1.40E-12	43	24	(Y17554) ornithine carbamoyltransferase [Bacillus licheniformis]
1760	Bt1Gc4535	Bt1G4167	1-206	g2293302	130	137	9.40E-09	45	21	(AF008220) YtqA [Bacillus subtilis]
1761	Bt1Gc4546	Bt1G4168	1-354	g2293279	136	195	1.60E-15	38	76	(AF008220) YtcG [Bacillus subtilis]
1762	Bt1Gc4534	Bt1G4169	1-312	g2293302	237	247	5.10E-21	51	32	(AF008220) YtqA [Bacillus subtilis]
1763	Bt1Gc4549	Bt1G4170	257-1	g3913544	187	184	1.00E-13	42	22	1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE (DXP REDUCTOISOMERASE) [Bacillus subtilis]
1764	Bt1Gc4551	Bt1G4171	268-1	g133481	172	203	2.30E-16	50	71	POSSIBLE RNA POLYMERASE SIGMA-G FACTOR (ORF3) [Bacillus thuringiensis]
1765	Bt1Gc4552	Bt1G4172	1-311	g120577	137	174	1.20E-12	40	26	CELL DIVISION PROTEIN FTSZ [Bacillus subtilis]
1766	Bt1Gc4554	Bt1G4173	1-409	g2497392	280	337	1.50E-30	51	55	INSERTION SEQUENCE IS232 PUTATIVE ATP-BINDING PROTEIN [Insertion sequence IS232]
1767	Bt1Gc4555	Bt1G4174	1-372	g2982194	108	131	1.50E-06	26	2	(AF007865) bacitracin synthetase 1; BacA [Bacillus licheniformis]
1768	Bt1Gc4556	Bt1G4175	1-400	g1168621	269	307	2.20E-27	52	43	BACITRACIN TRANSPORT ATP-BINDING PROTEIN BCRA [Bacillus licheniformis]
1769	Bt1Gc4558	Bt1G4176	1-274	g2500973	256	272	2.40E-23	53	21	ASPARTYL-TRNA SYNTHETASE (ASPARTATE--TRNA LIGASE) (ASPRS) [Pyrococcus sp.]
1770	Bt1Gc4573	Bt1G4177	1-235	g2634028	103	130	9.40E-08	39	18	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
1771	Bt1Gc4593	Bt1G4178	1-235	g3915700	130	102	6.90E-05	44	23	GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (GAPDH) [Bacillus

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1772	Bt1Gc4596	Bt1G4179	1-148	g1934659	80	102	6.50E-05	46	15	steaerothermophilus] (U93876) hypothetical protein YrdR [Bacillus subtilis]
1773	Bt1Gc4598	Bt1G4180	125-337	g2522020	112	149	1.20E-10	41	99	(AB007638) function unknown [Bacillus subtilis]
1774	Bt1Gc4610	Bt1G4181	1-286	g2635226	124	152	1.00E-09	41	12	(Z99118) similar to single-strand DNA-specific exonuclease [Bacillus subtilis]
1775	Bt1Gc4644	Bt1G4182	1-186	g127210	107	85	0.0085	44	12	METHYLMALONATE-SEMIALDEHYDE DEHYDROGENASE (ACYLATING) (MMSDH) [Pseudomonas aeruginosa]
1776	Bt1Gc4657	Bt1G4183	1-247	g1730898	165	195	2.90E-15	48	25	HYPOTHETICAL 36.3 KD PROTEIN IN RECQ-CMK INTERGENIC REGION [Bacillus subtilis]
1777	Bt1Gc4659	Bt1G4184	1-459	g1945117	287	269	2.40E-23	40	62	(D88802) transmembrane [Bacillus subtilis]
1778	Bt1Gc4670	Bt1G4185	1-334	g2506131	171	209	1.00E-15	45	12	ACONITATE HYDRATASE (CITRATE HYDRO-LYASE) (ACONITASE) [Bacillus subtilis]
1779	Bt1Gc4674	Bt1G4186	1-258	g1176702	152	187	1.20E-14	48	75	HYPOTHETICAL 13.0 KD PROTEIN IN QCRC-DAPB INTERGENIC REGION [Bacillus subtilis]
1780	Bt1Gc4673	Bt1G4187	271-1	g2589195	202	244	4.90E-20	60	18	(AF008553) Glu-tRNA <sup>Gln</sup> amidotransferase subunit A [Bacillus subtilis]
1781	Bt1Gc4679	Bt1G4188	1-429	g3123227	379	396	8.30E-37	56	27	GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE) [Bacillus subtilis]
1782	Bt1Gc4681	Bt1G4189	1-253	g2462120	317	319	1.20E-28	72	37	(Y11140) homology to M. leprae B2235_C2_195 and H. influenzae GB:U00019_14 [Bacillus cereus]
1783	Bt1Gc4687	Bt1G4190	472-1	g2497534	181	200	3.10E-15	32	33	PYRUVATE KINASE (PK) [Synecocystis sp.]
1784	Bt1Gc4691	Bt1G4191	377-18	g729484	139	217	9.00E-18	39	36	IRON-UPTAKE SYSTEM PROTEIN FEUB [Bacillus subtilis]
1785	Bt1Gc4695	Bt1G4192	291-1	g4481749	206	227	6.70E-19	49	40	(AF007865) BacR [Bacillus licheniformis]
1786	Bt1Gc4690	Bt1G4193	453-1	g2116974	220	282	2.00E-23	43	14	(D87979) YfnJ [Bacillus subtilis]
1787	Bt1Gc4701	Bt1G4194	148-1	g586883	158	163	4.10E-12	70	28	STAGE V SPOULATION PROTEIN T [Bacillus subtilis]
1788	Bt1Gc4703	Bt1G4195	523-226	g595780	122	110	1.90E-06	45	52	(U13871) lacZ alpha peptide [Cloning vector]
1789	Bt1Gc4714	Bt1G4196	4-371	g3288585	115	118	3.00E-06	34	21	(AJ224964) preprotein translocase [Phormidium lamosum]
1790	Bt1Gc4719	Bt1G4197	1-270	g1770029	191	209	5.40E-17	50	35	(Z75208) hypothetical protein [Bacillus subtilis]
1791	Bt1Gc4737	Bt1G4198	1-373	g128405	102	145	3.30E-10	31	57	NODULIN 21 (N-21) [Glycine max]
1792	Bt1Gc4746	Bt1G4199	1-287	g231698	125	132	6.30E-08	38	21	CATALASE [ ]
1793	Bt1Gc4750	Bt1G4200	1-240	g2293164	231	236	1.80E-19	60	20	(AF008220) SAM synthase [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvr	NCBI gi description
1794	Bt1Gc4730	Bt1G4201	1-249	g2293207	110	123	9.70E-08	36	38	(AF008220) YtmQ [Bacillus subtilis]
1795	Bt1Gc4759	Bt1G4202	131-423	g1196998	220	208	2.50E-16	55	24	(J01829) unknown protein [Transposon Tn10]
1796	Bt1Gc4763	Bt1G4203	336-87	g2633696	72	111	2.40E-05	30	11	(Z99110) similar to hypothetical proteins [Bacillus subtilis]
1797	Bt1Gc4773	Bt1G4204	254-31	g1196998	160	180	3.00E-13	51	18	(J01829) unknown protein [Transposon Tn10]
1798	Bt1Gc4779	Bt1G4205	453-1	g2661433	131	182	9.60E-13	31	14	(AJ001805) carbonylphosphate synthetase large subunit [Bacillus stearothermophilus]
1799	Bt1Gc4781	Bt1G4206	185-1	g2129564	80	110	2.30E-05	37	10	cryptochrome 2 apoprotein - Arabidopsis thaliana []
1800	Bt1Gc4791	Bt1G4207	1-470	g399784	372	372	2.90E-34	58	36	GLUTAMATE-1-SEMIALDEHYDE 2,1-AMINOMUTASE (GSA) (GLUTAMATE-1-SEMIALDEHYDE AMINOTRANSFERASE) (GSA-AT) [Bacillus subtilis]
1801	Bt1Gc4795	Bt1G4208	1-208	g2340005	246	265	6.30E-23	76	43	(Z98682) YibI protein [Bacillus subtilis]
1802	Bt1Gc4798	Bt1G4209	1-460	g3323354	134	164	3.50E-11	29	26	(AE001270) oligoendopeptidase F, putative [Treponema pallidum]
1803	Bt1Gc4804	Bt1G4210	155-480	g1648861	199	222	9.20E-18	44	25	(Z81356) UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Bacillus subtilis]
1804	Bt1Gc4802	Bt1G4211	1-287	g2497382	213	248	1.10E-20	56	22	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1805	Bt1Gc4808	Bt1G4212	228-39	g1769947	109	131	2.70E-07	47	6	(X98455) SNF [Bacillus cereus]
1806	Bt1Gc4809	Bt1G4213	1-449	g1705442	322	316	2.50E-28	42	41	PUTATIVE BRANCHED-CHAIN-FATTY-ACID KINASE [Bacillus subtilis]
1807	Bt1Gc4851	Bt1G4214	1-259	g1934835	119	142	8.60E-09	34	14	(Z93940) asparagine synthetase [Bacillus subtilis]
1808	Bt1Gc4855	Bt1G4215	1-290	g2632519	173	210	4.70E-17	42	30	(Z99105) similar to hypothetical proteins [Bacillus subtilis]
1809	Bt1Gc4856	Bt1G4216	1-213	g114510	124	142	6.30E-09	46	14	ATP SYNTHASE ALPHA CHAIN [Bacillus megaterium]
1810	Bt1Gc4869	Bt1G4217	1-415	g2494075	126	161	5.50E-11	29	28	NADP-DEPENDENT GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (NON-PHOSPHORYLATING GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE) (GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (NADP+)) (TRIOSEPHOSPHATE DEHYDROGENASE) [Zea mays]
1811	Bt1Gc4862	Bt1G4218	205-2	g584768	145	122	3.40E-07	47	23	N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE (AGPR) (N-ACETYL-GLUTAMATE SEMIALDEHYDE DEHYDROGENASE) (NAGSA DEHYDROGENASE) [Bacillus stearothermophilus]
1812	Bt1Gc4876	Bt1G4219	406-1	g2635847	196	192	2.10E-14	37	29	(Z99121) similar to ABC transporter (amino acid permease)

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1813	Bt1Gc4881	Bt1G4220	1-224	g1695686	133	115	1.50E-05	48	6	[Bacillus subtilis] (D83706) pyruvate carboxylase [Bacillus stearothermophilus]
1814	Bt1Gc4884	Bt1G4221	1-169	g225559	169	173	2.60E-12	63	12	ORF IS231C [Bacillus thuringiensis]
1815	Bt1Gc4890	Bt1G4222	245-1	g4033397	145	183	5.40E-13	43	10	DNA GYRASE SUBUNIT A [Bacillus sp.]
1816	Bt1Gc4894	Bt1G4223	384-1	g2501371	170	184	2.40E-14	36	46	SULFATE TRANSPORT SYSTEM PERMEASE PROTEIN CYSW [Synecocystis sp.]
1817	Bt1Gc4895	Bt1G4224	1-377	g2116767	231	273	7.60E-23	40	19	(D86418) YfnI [Bacillus subtilis]
1818	Bt1Gc4903	Bt1G4225	1-458	g2116767	388	332	3.00E-29	52	23	(D86418) YfnI [Bacillus subtilis]
1819	Bt1Gc4905	Bt1G4226	1-227	g2635181	115	202	3.10E-16	60	24	(Z99117) similar to protease [Bacillus subtilis]
1820	Bt1Gc4914	Bt1G4227	1-262	g1169919	145	149	1.50E-09	38	15	GLUCOSAMINE--FRUCTOSE-6-PHOSPHATE AMINOTRANSFERASE (ISOMERIZING) (HEXOSEPHOSPHATE AMINOTRANSFERASE) (D- FRUCTOSE-6-PHOSPHATE AMIDOTRANSFERASE) (GFAT) (L-GLUTAMINE-D-FRUCTOSE-6-PHOSPHATE AMIDOTRANSFERASE) (GLUCOSAMINE-6-PH...
1821	Bt1Gc4920	Bt1G4228	1-186	g2988346	112	156	2.00E-10	49	12	[Bacillus subtilis] (D89812) catalase [Sus scrofa]
1822	Bt1Gc4921	Bt1G4229	203-1	g1124822	182	212	1.50E-16	66	14	(M16158) unknown protein [Bacillus thuringiensis]
1823	Bt1Gc4954	Bt1G4230	420-21	g2127273	679	681	5.20E-67	98	83	B. subtilis spoIIIGA protein homolog - Bacillus thuringiensis II
1824	Bt1Gc4956	Bt1G4231	1-189	g732384	96	106	2.70E-05	40	18	HYPOTHETICAL 37.0 KD PROTEIN IN SPOIIR-GLYC INTERGENIC REGION [Bacillus subtilis]
1825	Bt1Gc4962	Bt1G4232	160-1	g790866	95	113	2.70E-06	46	20	(L36381) orfB; putative [Neisseria gonorrhoeae]
1826	Bt1Gc4997	Bt1G4233	20-379	g3328184	225	235	9.50E-20	43	42	(AF074376) beta lactamase precursor [Cloning vector pQE30NST]
1827	Bt1Gc5002	Bt1G4234	206-1	g595780	104	92	0.00025	47	37	(U13871) lacZ alpha peptide [Cloning vector]
1828	Bt1Gc5004	Bt1G4235	1-388	g2635600	98	128	1.30E-07	19	33	(Z99119) similar to hypothetical proteins [Bacillus subtilis]
1829	Bt1Gc5003	Bt1G4236	1-461	g2635858	207	259	3.50E-21	34	20	(Z99121) yvgS [Bacillus subtilis]
1830	Bt1Gc5005	Bt1G4237	383-1	g1945657	138	160	1.10E-11	33	49	(Z94043) hypothetical protein [Bacillus subtilis]
1831	Bt1Gc5021	Bt1G4238	1-306	g2116974	224	263	2.20E-21	54	10	(D87979) YfnJ [Bacillus subtilis]
1832	Bt1Gc5027	Bt1G4239	279-1	g732355	107	135	8.20E-09	33	36	HYPOTHETICAL 28.4 KD PROTEIN IN SACT-SACP INTERGENIC REGION [Bacillus subtilis]
1833	Bt1Gc5039	Bt1G4240	327-1	g3123297	203	233	1.60E-18	54	16	EXCINUCLEASE ABC SUBUNIT B (DINA PROTEIN) [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvr	NCBI gi description
1834	Bt1Gc5047	Bt1G4241	1-358	g1619837	141	67	0.00064	39	39	(U67921) collagen-like protein [Bacillus thuringiensis israelensis]
1835	Bt1Gc5056	Bt1G4242	1-421	g2633541	184	240	4.00E-20	40	37	(Z99110) similar to cystathionine gamma-synthase [Bacillus subtilis]
1836	Bt1Gc5065	Bt1G4243	57-296	g2577965	186	197	1.40E-14	48	11	(Y15254) PcrA protein [Bacillus subtilis]
1837	Bt1Gc5069	Bt1G4244	314-63	g729341	98	138	1.10E-08	34	21	DLTB PROTEIN [Bacillus subtilis]
1838	Bt1Gc5070	Bt1G4245	400-26	g225559	277	219	2.60E-17	48	26	ORF IS231C [Bacillus thuringiensis]
1839	Bt1Gc5073	Bt1G4246	114-522	g2619012	411	436	4.80E-41	61	39	(AF027868) fatty acid desaturase [Bacillus subtilis]
1840	Bt1Gc5081	Bt1G4247	34-321	g266569	171	150	2.50E-10	41	32	ROD SHAPE-DETERMINING PROTEIN MREC [Bacillus stearothermophilus]
1841	Bt1Gc5080	Bt1G4248	1-379	g730399	322	344	1.30E-30	58	20	PRKA PROTEIN [Bacillus subtilis]
1842	Bt1Gc5087	Bt1G4249	489-1	g3025180	524	524	2.30E-50	60	36	HYPOTHETICAL 49.9 KD PROTEIN IN CITA-SSPB INTERGENIC REGION [Bacillus subtilis]
1843	Bt1Gc5098	Bt1G4250	1-369	g729340	136	220	2.30E-17	35	24	D-ALANINE-ACTIVATING ENZYME (DAE) (D-ALANINE-D-ALANYL CARRIER PROTEIN LIGASE) (DCL) [Bacillus subtilis]
1844	Bt1Gc5114	Bt1G4251	1-446	g1731364	194	242	3.70E-20	37	37	HYPOTHETICAL PROTEASE IN ROCR-PURA INTERGENIC REGION [Bacillus subtilis]
1845	Bt1Gc5130	Bt1G4252	1-395	g2635723	87	127	2.10E-07	26	30	(Z99120) homoserine dehydrogenase [Bacillus subtilis]
1846	Bt1Gc5131	Bt1G4253	1-256	g1729789	111	147	1.50E-09	34	19	(Y09927) phosphoglucosamine mutase [Staphylococcus aureus]
1847	Bt1Gc5144	Bt1G4254	374-1	g585920	326	363	5.50E-32	62	10	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Bacillus subtilis]
1848	Bt1Gc5150	Bt1G4255	1-386	g1176955	212	234	2.90E-19	40	32	HYPOTHETICAL 45.8 KD PROTEIN IN ACDA-NARI INTERGENIC REGION [Bacillus subtilis]
1849	Bt1Gc5159	Bt1G4256	67-362	g2226151	182	212	9.40E-17	43	24	(Y14080) hypothetical protein [Bacillus subtilis]
1850	Bt1Gc5161	Bt1G4257	1-139	g2144427	89	113	1.30E-05	40	7	phosphotransferase system enzyme II (EC 2.7.1.69), glucose-specific, factor IIB - Staphylococcus carnosus [Staphylococcus carnosus]
1851	Bt1Gc5162	Bt1G4258	1-386	g2462090	97	130	1.20E-07	30	25	(Y11171) BC542A protein [Bacillus cereus]
1852	Bt1Gc5166	Bt1G4259	220-1	g2635763	117	140	9.20E-09	44	16	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
1853	Bt1Gc5171	Bt1G4260	1-300	g2443243	87	110	1.70E-06	30	85	(D86417) YfIT [Bacillus subtilis]
1854	Bt1Gc5183	Bt1G4261	450-1	g1945096	446	406	7.20E-38	59	35	(D88802) S. lividans chloramphenicol resistance protein; P31141 (492) transmembrane [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1855	Bt1Gc5212	Bt1G4262	297-1	g3122295	193	197	1.50E-15	49	31	6-PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) [Bacillus subtilis]
1856	Bt1Gc5214	Bt1G4263	1-302	g548931	238	192	1.40E-14	60	26	S-LAYER PROTEIN PRECURSOR (SURFACE LAYER PROTEIN) [Bacillus thuringiensis]
1857	Bt1Gc5225	Bt1G4264	533-1	g2293257	593	561	2.70E-54	66	60	(AF008220) YtmM [Bacillus subtilis]
1858	Bt1Gc5227	Bt1G4265	1-146	g448837	145	164	3.20E-12	57	23	superoxide dismutase [Bacillus stearothermophilus]
1859	Bt1Gc5246	Bt1G4266	84-376	g1709415	159	186	6.20E-14	39	25	PYRIMIDINE NUCLEOSIDE TRANSPORT PROTEIN [Bacillus subtilis]
1860	Bt1Gc5247	Bt1G4267	16-281	g1168281	97	153	2.40E-10	40	23	ACYL-COA DEHYDROGENASE [Bacillus subtilis]
1861	Bt1Gc5253	Bt1G4268	379-1	g3183483	319	331	6.40E-30	57	40	HYPOTHETICAL PROTEIN (ORF3) [Bacillus cereus]
1862	Bt1Gc5284	Bt1G4270	274-1	g1170977	165	129	1.50E-07	41	19	PROBABLE METHYLMALONATE-SEMIALDEHYDE DEHYDROGENASE [Bacillus subtilis]
1863	Bt1Gc5302	Bt1G4271	450-127	g3023940	95	103	9.30E-06	28	74	HIT PROTEIN [Bacillus subtilis]
1864	Bt1Gc5292	Bt1G4272	511-1	g113345	145	120	4.70E-06	34	12	ATP-DEPENDENT NUCLEASE SUBUNIT A [Bacillus subtilis]
1865	Bt1Gc5309	Bt1G4273	52-293	g4894251	144	148	1.30E-09	54	17	(AF065404) pXO1-35 [Bacillus anthracis]
1866	Bt1Gc5328	Bt1G4274	1-248	g321919	70	117	3.00E-07	44	42	hypothetical 16.9K protein - Salmonella typhimurium plasmid NTP16 []
1867	Bt1Gc5329	Bt1G4275	527-1	g126053	262	292	8.70E-26	36	55	L-LACTATE DEHYDROGENASE P [Bacillus psychrosaccharolyticus]
1868	Bt1Gc5333	Bt1G4276	1-229	g3122218	177	214	1.60E-17	57	24	HPR(SER) KINASE [Bacillus subtilis]
1869	Bt1Gc5339	Bt1G4277	383-1	g549114	93	161	5.10E-11	28	26	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS231F [Bacillus thuringiensis]
1870	Bt1Gc5340	Bt1G4278	1-322	g3915759	101	114	1.90E-05	19	9	MNN4 PROTEIN [Saccharomyces cerevisiae]
1871	Bt1Gc5341	Bt1G4279	1-238	g133083	166	184	2.40E-14	51	67	50S RIBOSOMAL PROTEIN L7/L12 (MA1/MA2) []
1872	Bt1Gc5342	Bt1G4280	1-180	g1173064	121	127	2.60E-08	48	48	50S RIBOSOMAL PROTEIN L7/L12 (BL9) (A' TYPE) (VEGETATIVE PROTEIN 341) (VEG341) [Bacillus subtilis]
1873	Bt1Gc5359	Bt1G4281	99-457	g2462088	170	187	1.20E-14	37	62	(Y11170) BC541A protein [Bacillus cereus]
1874	Bt1Gc5360	Bt1G4282	1-184	g114646	142	162	1.00E-11	55	21	ATP SYNTHASE GAMMA CHAIN PRECURSOR [thermophilic bacterium PS3]
1875	Bt1Gc5369	Bt1G4284	1-205	g987050	103	125	5.10E-08	36	33	(X65335) lacZ [Cloning vector pSV-beta-Galactosidase Control]
1876	Bt1Gc5388	Bt1G4285	469-1	g121881	340	351	5.00E-31	46	19	DNA GYRASE SUBUNIT A [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1877	Bt1Gc5386	Bt1G4286	490-130	g2633738	203	203	2.30E-16	46	86	(Z99111) similar to transcriptional regulator (MarR family) [Bacillus subtilis]
1878	Bt1Gc5406	Bt1G4287	382-1	g1731075	138	151	3.00E-10	29	38	PROBABLE NADH-DEPENDENT FLAVIN OXIDOREDUCTASE YQM [Bacillus subtilis]
1879	Bt1Gc5411	Bt1G4288	1-419	g2492737	187	228	8.80E-18	37	16	ALCOHOL DEHYDROGENASE 2 (ADH) / ALCTALDEHYDE DEHYDROGENASE (ACDH) [Entamoeba histolytica]
1880	Bt1Gc5412	Bt1G4289	388-1	g2618834	171	175	1.60E-12	28	27	(AF017113) putative protease [Bacillus subtilis]
1881	Bt1Gc5442	Bt1G4290	366-1	g2635670	143	183	2.20E-13	33	22	(Z99120) similar to nicotinate phosphoribosyltransferase [Bacillus subtilis]
1882	Bt1Gc5453	Bt1G4291	596-295	g2633170	244	219	6.50E-18	47	29	(Z99108) similar to iron(III) dicitrate transport permease [Bacillus subtilis]
1883	Bt1Gc5464	Bt1G4292	159-1	g1009035	184	184	2.30E-13	62	9	(L47692) pyruvate oxidase [Escherichia coli]
1884	Bt1Gc5466	Bt1G4293	1-144	g586858	119	129	1.60E-08	60	24	HYPOTHETICAL 21.4 KD PROTEIN IN DACA-SERS INTERGENIC REGION [Bacillus subtilis]
1885	Bt1Gc5472	Bt1G4294	1-388	g729583	66	130	6.30E-08	21	38	GLYCOGEN BIOSYNTHESIS PROTEIN GLGD [Bacillus subtilis]
1886	Bt1Gc5473	Bt1G4295	278-1	g2501426	255	301	9.70E-27	62	21	PYRIMIDINE-NUCLEOSIDE PHOSPHORYLASE [Bacillus stearothermophilus]
1887	Bt1Gc5454	Bt1G4296	269-1	g1945050	117	124	4.60E-07	32	20	(U63928) L2 protein [Bacillus cereus]
1888	Bt1Gc5467	Bt1G4297	1-255	g729582	109	130	7.80E-08	38	22	GLUCOSE-1-PHOSPHATE ADENYL-TRANSFERASE (ADP-GLUCOSE SYNTHASE) (ADP-GLUCOSE PYROPHOSPHORYLASE) [Bacillus subtilis]
1889	Bt1Gc5486	Bt1G4298	1-152	g2634032	102	128	1.20E-07	53	13	(Z99112) nusA [Bacillus subtilis]
1890	Bt1Gc5488	Bt1G4299	465-1	g225559	458	477	2.20E-45	63	32	ORF IS231C [Bacillus thuringiensis]
1891	Bt1Gc5498	Bt1G4300	447-1	g3123286	144	161	1.20E-10	31	18	LEUCYL-TRNA SYNTHETASE (LEUCINE--TRNA LIGASE) (LEURS) [Bacillus subtilis]
1892	Bt1Gc5501	Bt1G4301	600-1	g586848	117	211	7.90E-17	34	53	HYPOTHETICAL 40.1 KD GTP-BINDING PROTEIN IN RPSF-SPOOJ INTERGENIC REGION [Bacillus subtilis]
1893	Bt1Gc5521	Bt1G4302	420-90	g729325	108	143	1.10E-09	32	42	GLUCOSE 1-DEHYDROGENASE II (GLCDH-II) [Bacillus megaterium]
1894	Bt1Gc5527	Bt1G4303	306-1	g730776	200	220	5.40E-17	49	13	STAGE II SPORULATION PROTEIN E [Bacillus subtilis]
1895	Bt1Gc5536	Bt1G4304	322-1	g1750115	105	129	1.60E-08	36	62	(U66480) YnaD [Bacillus subtilis]
1896	Bt1Gc5563	Bt1G4306	1-449	g3914433	115	203	1.40E-15	42	19	PRISMANE PROTEIN HOMOLOG [Methanobacterium thermoautotrophicum]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1897	Bt1Gc5568	Bt1G4307	356-1	g118793	302	292	2.70E-24	55	8	DNA POLYMERASE III, ALPHA CHAIN [Bacillus subtilis]
1898	Bt1Gc5574	Bt1G4308	190-1	g3777577	127	146	5.60E-10	46	23	(AF083408) kanamycin 3'-phosphotransferase [Cloning vector p34S-Km3]
1899	Bt1Gc5585	Bt1G4309	186-1	g2462104	145	152	7.10E-11	52	25	(Y10981) gamma-glutamyl phosphate reductase [Bacillus cereus]
1900	Bt1Gc5590	Bt1G4310	148-5	g121415	119	121	1.10E-06	46	9	GLYCEROL KINASE (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE) (GLYCEROKINASE) (GK) [Bacillus subtilis]
1901	Bt1Gc5592	Bt1G4311	1-188	g118672	125	138	1.50E-08	46	13	LIPOAMIDE DEHYDROGENASE COMPONENT OF PYRUVATE DEHYDROGENASE COMPLEX (E3) (DIHYDROLIPOAMIDE DEHYDROGENASE) (S COMPLEX, 50 KD SUBUNIT) [Bacillus subtilis]
1902	Bt1Gc5609	Bt1G4312	1-496	g1789981	258	286	3.80E-25	35	58	(AE000433) IS150 putative transposase [Escherichia coli]
1903	Bt1Gc5627	Bt1G4313	171-1	g121887	92	114	9.20E-06	40	9	DNA GYRASE SUBUNIT B [Bacillus subtilis]
1904	Bt1Gc5637	Bt1G4314	1-150	g1709188	86	103	0.00014	43	8	DNA MISMATCH REPAIR PROTEIN MUTL [Bacillus subtilis]
1905	Bt1Gc5636	Bt1G4315	1-199	g2522016	85	106	2.70E-05	41	19	(AB007638) dehydrogenase [Bacillus subtilis]
1906	Bt1Gc5642	Bt1G4317	22-481	g115022	380	353	3.00E-32	58	59	BETA-LACTAMASE PRECURSOR, TYPE II (PENICILLINASE) (CEPHALOSPORINASE) [Bacillus cereus]
1907	Bt1Gc5647	Bt1G4318	1-415	g117576	170	62	0.79	30	41	CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS) [Plasmodium berghei]
1908	Bt1Gc5656	Bt1G4320	1-409	g4584100	222	215	1.30E-17	43	81	(AJ010132) hypothetical protein [Bacillus cereus]
1909	Bt1Gc5662	Bt1G4321	319-1	g4033396	300	321	4.40E-28	61	17	DNA GYRASE SUBUNIT B [Bacillus sp.]
1910	Bt1Gc5671	Bt1G4322	1-421	g3123300	275	261	1.70E-22	49	44	THIOREDOXIN REDUCTASE (GENERAL STRESS PROTEIN 35) (GSP35) [Bacillus subtilis]
1910	Bt1Gc5671	Bt1G4323	1-421	g4379428	231	227	6.70E-19	46	45	(AJ223781) thioredoxin reductase [Staphylococcus aureus]
1911	Bt1Gc5675	Bt1G4324	161-1	g1524023	95	117	1.20E-06	49	19	(X95401) beta-lactamase [Escherichia coli]
1912	Bt1Gc5677	Bt1G4325	272-1	g2635723	153	83	0.0018	43	21	(Z99120) homoserine dehydrogenase [Bacillus subtilis]
1913	Bt1Gc5681	Bt1G4326	282-1	g1172710	158	170	4.90E-12	40	21	PTS SYSTEM, CELLOBIOSE-SPECIFIC IIC COMPONENT (EII-C-CEL) (CELLOBIOSE-PERMEASE IIC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, C COMPONENT) [Bacillus subtilis]
1914	Bt1Gc5686	Bt1G4327	1-275	g2668553	130	68	0.85	37	6	(U62929) multidrug resistance protein 1 [Filobasidiella

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1915	Bt1Gc5715	Bt1G4330	282-1	g4512436	84	112	3.40E-06	28	35	neoformans] (AB017508) trnA homologue (identity of 62% to B. subtilis) [Bacillus halodurans]
1916	Bt1Gc5727	Bt1G4332	380-125	g1706300	197	216	9.80E-18	53	32	DIHYDROPICOLINATE REDUCTASE [Bacillus subtilis]
1917	Bt1Gc5732	Bt1G4333	1-240	g3318590	141	121	1.10E-07	40	45	(AB015670) A2-5aorf21; hypothetical protein [Bacillus sp.]
1918	Bt1Gc5748	Bt1G4334	411-1	g124423	204	92	0.0016	37	27	INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE (IMP DEHYDROGENASE) (IMPDH) (IMPD) (SUPEROXIDE-INDUCIBLE PROTEIN 12) (SOI12) [Bacillus subtilis]
1919	Bt1Gc5751	Bt1G4335	212-1	g586893	131	136	4.20E-09	38	31	HYPOTHETICAL 26.2 KD PROTEIN IN FTSH-CYSK INTERGENIC REGION [Bacillus subtilis]
1920	Bt1Gc5761	Bt1G4336	1-205	g225559	153	55	0.0035	54	14	ORF IS231C [Bacillus thuringiensis]
1921	Bt1Gc5771	Bt1G4337	1-417	g1171068	256	337	1.50E-30	51	32	PROBABLE UDP-N-ACETYLGLUCOSAMINE 1-CARBOXYVINYLTRANSFERASE (ENOYLPIRUVATE TRANSFERASE) (UDP-N-ACETYLGLUCOSAMINE ENOLPYRUVYL TRANSFERASE) (EPT) [Bacillus subtilis]
1922	Bt1Gc5768	Bt1G4338	1-500	g4512389	331	185	2.40E-13	45	25	(AB011838) methyl-accepting chemotaxis protein [Bacillus halodurans]
1923	Bt1Gc5767	Bt1G4339	254-70	g115950	133	153	1.80E-10	55	18	GLUCOSE-RESISTANCE AMYLASE REGULATOR (CATABOLITE CONTROL PROTEIN) [Bacillus subtilis]
1924	Bt1Gc5772	Bt1G4340	319-1	g143434	276	282	1.30E-24	59	25	(M97678) Rho Factor [Bacillus subtilis]
1925	Bt1Gc5779	Bt1G4341	1-459	g1834395	276	333	3.90E-30	52	55	(X67138) flagellin [Bacillus thuringiensis]
1926	Bt1Gc5777	Bt1G4342	263-1	g461914	77	135	2.20E-08	35	21	PENICILLIN-BINDING PROTEIN 5* PRECURSOR (D-ALANYL-D-ALANINE CARBOXYPEPTIDASE) (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) (PBP-5*) [Bacillus subtilis]
1927	Bt1Gc5796	Bt1G4344	1-262	g2635774	228	191	4.40E-15	51	59	(Z99120) yusF [Bacillus subtilis]
1927	Bt1Gc5796	Bt1G4345	21-553	g2635773	198	229	4.10E-19	45	91	(Z99120) similar to thioredoxin [Bacillus subtilis]
1928	Bt1Gc5802	Bt1G4346	1-182	g1350849	117	133	1.80E-07	50	5	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Staphylococcus aureus]
1929	Bt1Gc5808	Bt1G4347	1-244	g3122229	176	188	2.40E-14	51	24	HEAT-INDUCIBLE TRANSCRIPTION REPRESSOR

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat <sub>nap</sub> Score	BlastP Score	BlastP Prob	% Ident	% Cvr	NCBI gi description
1930	Bt1Gc5818	Bt1G4348	1-365	g584719	213	209	1.80E-16	45	30	HRCA [Bacillus stearothermophilus]
1931	Bt1Gc5837	Bt1G4349	475-1	g541317	71	113	1.80E-05	35	8	ACETATE KINASE (ACETOKINASE) [Bacillus subtilis] H <sup>+</sup> -transporting ATPase (EC 3.6.1.35) PMA1 - Synecocystis sp [Synecocystis sp.]
1932	Bt1Gc5838	Bt1G4350	345-1	g2415403	238	213	9.00E-17	44	26	(AF015775) acetylornithine deacetylase [Bacillus subtilis]
1933	Bt1Gc5835	Bt1G4351	1-368	g417116	190	228	5.30E-19	38	40	FERROCHELATASE (PROTOHEME FERRO-LYASE) (HEME SYNTHETASE) [Bacillus subtilis]
1934	Bt1Gc5841	Bt1G4352	406-1	g417116	111	147	6.60E-10	28	44	FERROCHELATASE (PROTOHEME FERRO-LYASE) (HEME SYNTHETASE) [Bacillus subtilis]
1935	Bt1Gc5877	Bt1G4353	1-151	g2127147	152	181	3.60E-13	63	10	nitrate reductase (EC 1.7.99.4) beta chain - Bacillus subtilis
1936	Bt1Gc5882	Bt1G4354	1-395	g3023410	337	378	6.70E-35	63	27	BRANCHED-CHAIN AMINO ACID TRANSPORT SYSTEM CARRIER PROTEIN (BRANCHED-CHAIN AMINO ACID UPTAKE CARRIER) [Bacillus subtilis]
1937	Bt1Gc5889	Bt1G4355	377-1	g1805463	198	253	1.20E-21	46	61	(D50453) ycnI [Bacillus subtilis]
1938	Bt1Gc5899	Bt1G4357	98-268	g4589971	265	302	7.60E-27	86	58	(AC007195) putative blue copper-binding protein II [Arabidopsis thaliana]
1939	Bt1Gc5903	Bt1G4359	115-318	g2635515	115	118	2.40E-07	38	99	(Z99119) similar to hypothetical proteins from B. subtilis [Bacillus subtilis]
1940	Bt1Gc5917	Bt1G4360	298-1	g585047	189	233	2.90E-19	54	26	ALANINE DEHYDROGENASE (STAGE V SPORULATION PROTEIN N) [Bacillus subtilis]
1941	Bt1Gc5918	Bt1G4361	1-301	g729934	266	277	3.40E-24	50	55	SIGNAL PEPTIDASE I (SPASE I) (LEADER PEPTIDASE I) [Bacillus caldolyticus]
1941	Bt1Gc5918	Bt1G4362	325-471	g4514314	223	223	1.80E-18	86	17	(AB013365) YlqF [Bacillus halodurans]
1942	Bt1Gc5913	Bt1G4363	520-191	g4309751	415	447	3.30E-42	75	27	(AC006217) hypothetical protein [Arabidopsis thaliana]
1943	Bt1Gc5939	Bt1G4364	387-1	g2506131	314	373	2.70E-33	62	14	ACONITATE HYDRATASE (CITRATE HYDRO-LYASE) (ACONITASE) [Bacillus subtilis]
1944	Bt1Gc5942	Bt1G4365	45-325	g1255196	269	199	6.00E-15	53	16	(U50744) BSMA [Bacillus stearothermophilus]
1945	Bt1Gc5948	Bt1G4366	1-355	g225559	246	204	1.10E-15	50	25	ORF IS231C [Bacillus thuringiensis]
1946	Bt1Gc5949	Bt1G4367	355-1	g2635847	292	310	1.50E-27	45	25	(Z99121) similar to ABC transporter (amino acid permease) [Bacillus subtilis]
1947	Bt1Gc5937	Bt1G4368	1-437	g1162917	213	241	2.20E-20	41	60	(L48554) vrrA gene product [Bacillus anthracis]
1948	Bt1Gc5953	Bt1G4369	1-333	g225559	296	322	7.10E-29	63	23	ORF IS231C [Bacillus thuringiensis]
1949	Bt1Gc5968	Bt1G4370	186-1	g399058	108	135	2.10E-08	45	17	CHORISMATE SYNTHASE (5-ENOLPYRUVYL-SHIKIMATE-3-PHOSPHATE

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aaf_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
										PHOSPHOLYASE) (VEGETATIVE PROTEIN 216) (VEG216) [Bacillus subtilis]
1950	Bt1Gc5970	Bt1G4371	1-319	g1934614	110	159	2.70E-10	32	10	(U93874) cytochrome P450 102 [Bacillus subtilis]
1951	Bt1Gc5971	Bt1G4372	1-267	g1813466	88	118	4.40E-07	34	38	(U61168) spore germination protein C2 [Bacillus firmus]
1952	Bt1Gc5979	Bt1G4373	1-291	g131720	95	121	5.00E-07	31	31	DIHYDROOROTATE DEHYDROGENASE (DIHYDROOROTATE OXIDASE) (DHODEHASE) [Bacillus subtilis]
1953	Bt1Gc5992	Bt1G4374	1-358	g399058	216	264	8.00E-23	51	30	CHORISMATE SYNTHASE (5- ENOLPYRUVYL-SHIKIMATE-3-PHOSPHATE PHOSPHOLYASE) (VEGETATIVE PROTEIN 216) (VEG216) [Bacillus subtilis]
1954	Bt1Gc5990	Bt1G4375	450-1	g2492538	175	163	1.60E-11	37	41	SULFATE TRANSPORT ATP-BINDING PROTEIN CYSA [Synechocystis sp.]
1955	Bt1Gc5995	Bt1G4376	1-332	g3025120	137	172	5.50E-12	38	17	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YDIF [Bacillus subtilis]
1956	Bt1Gc5998	Bt1G4377	1-385	g2226226	124	198	4.60E-15	34	27	(Y14083) Hypothetical protein [Bacillus subtilis]
1957	Bt1Gc6014	Bt1G4378	1-307	g1706723	126	154	2.60E-10	32	23	PUTATIVE EXODEOXYRIBONUCLEASE LARGE SUBUNIT (EXONUCLEASE VII LARGE SUBUNIT) [Bacillus subtilis]
1958	Bt1Gc6030	Bt1G4379	19-441	g4512352	465	466	3.20E-44	60	28	(AB011836) proline transporter [Bacillus halodurans]
1959	Bt1Gc6035	Bt1G4380	477-66	g2829796	182	216	5.20E-17	33	29	CYTOCHROME D UBIQUINOL OXIDASE SUBUNIT I [Bacillus subtilis]
1960	Bt1Gc6038	Bt1G4381	217-1	g2522016	150	184	7.00E-14	43	21	(AB007638) dehydrogenase [Bacillus subtilis]
1961	Bt1Gc6047	Bt1G4382	246-1	g3183561	170	206	1.10E-16	52	31	VEGETATIVE PROTEIN 296 (VEG296) [Bacillus subtilis]
1962	Bt1Gc6046	Bt1G4383	264-1	g2618861	110	133	6.10E-09	38	51	(AF017113) putative acetyltransferase [Bacillus subtilis]
1963	Bt1Gc6053	Bt1G4384	1-471	g2632105	584	607	3.60E-59	74	24	(Z98760) arginyl-tRNA synthetase [Arabidopsis thaliana]
1964	Bt1Gc6055	Bt1G4385	450-1	g2577965	195	201	5.30E-15	34	20	(Y15254) PcrA protein [Bacillus subtilis]
1965	Bt1Gc6057	Bt1G4386	515-1	g2497382	218	203	1.10E-15	36	40	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232] [Bacillus subtilis]
1966	Bt1Gc6071	Bt1G4387	192-17	g2984721	169	171	5.80E-13	64	91	(AF053927) probable spore germination protein D [Bacillus cereus]
1967	Bt1Gc6069	Bt1G4388	553-1	g1175719	441	469	1.50E-44	46	28	PENICILLIN-BINDING PROTEIN 3 (PBP 3) [Bacillus subtilis]
1968	Bt1Gc6083	Bt1G4389	277-1	g586902	107	167	6.10E-12	38	26	HYPOTHETICAL 40.7 KD PROTEIN IN MECB-GLTX INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1969	Bt1Gc6086	Bt1G4390	1-371	g1350848	201	278	7.30E-23	51	9	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Spiroplasma citri]
1970	Bt1Gc6097	Bt1G4391	1-356	g2633807	158	174	2.80E-13	33	52	(Z99111) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis]
1971	Bt1Gc6103	Bt1G4392	1-321	g730959	95	118	3.60E-06	24	16	METHYL-ACCEPTING CHEMOTAXIS PROTEIN TLPB [Bacillus subtilis]
1972	Bt1Gc6108	Bt1G4393	1-277	g2529473	173	191	4.40E-15	43	34	(AF006665) YokZ [Bacillus subtilis]
1973	Bt1Gc6114	Bt1G4394	332-1	g124423	225	159	9.60E-11	50	22	INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE (IMP DEHYDROGENASE) (IMPDH) (IMPD) (SUPEROXIDE-INDUCIBLE PROTEIN 12) (SOI12) [Bacillus subtilis]
1974	Bt1Gc6119	Bt1G4395	1-233	g4557058	93	140	3.80E-08	58	4	(AC007154) aldehyde oxidase [Arabidopsis thaliana]
1975	Bt1Gc6143	Bt1G4396	1-334	g585481	258	283	1.90E-23	50	9	TRANSCRIPTION-REPAIR COUPLING FACTOR (TRCF) [Bacillus subtilis]
1976	Bt1Gc6164	Bt1G4397	1-367	g3915673	111	166	6.20E-11	32	10	PROBABLE SERINE ACTIVATING ENZYME [Bacillus subtilis]
1977	Bt1Gc6184	Bt1G4398	1-183	g2117766	105	117	2.20E-06	43	15	orthinine aminotransferase rocD - Bacillus subtilis []
1978	Bt1Gc6197	Bt1G4399	566-391	g2245063	270	188	1.50E-14	98	19	(Z97342) hypothetical protein [Arabidopsis thaliana]
1979	Bt1Gc6210	Bt1G4402	270-460	g1730905	112	125	4.30E-08	40	52	HYPOTHETICAL 13.7 KD PROTEIN IN BC5A-DEGR INTERGENIC REGION [Bacillus subtilis]
1980	Bt1Gc6215	Bt1G4403	1-366	g3913544	252	258	4.00E-22	43	31	1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE (DXP REDUCTOISOMERASE) [Bacillus subtilis]
1981	Bt1Gc6217	Bt1G4404	394-1	g68538	117	164	6.30E-11	33	15	valine-tRNA ligase (EC 6.1.1.9) - Bacillus stearothermophilus []
1982	Bt1Gc6231	Bt1G4405	1-177	g584768	112	110	7.20E-06	42	20	N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE (AGPR) (N-ACETYL-GLUTAMATE SEMIALDEHYDE DEHYDROGENASE) (NAGSA DEHYDROGENASE) [Bacillus stearothermophilus]
1983	Bt1Gc6239	Bt1G4406	1-529	g2982194	165	239	5.30E-18	28	3	(AF007865) bacitracin synthetase 1; BacA [Bacillus licheniformis]
1984	Bt1Gc6242	Bt1G4407	114-299	g2996158	232	180	6.40E-14	63	98	(AF051753) GAS5A-like protein [Picea mariana]
1985	Bt1Gc6248	Bt1G4408	1-218	g3395452	121	159	1.10E-11	43	30	(X92496) UbiH-like protein [Azospirillum brasilense]
1986	Bt1Gc6252	Bt1G4410	309-1	g3319360	162	205	3.60E-15	64	5	(AF077408) contains similarity to Vicia faba

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1987	Bt1Gc6265	Bt1G4411	1-482	g2342692	159	322	3.70E-28	66	14	retrotransposon-like gene (GB:AB007467) [Arabidopsis thaliana]
1988	Bt1Gc6274	Bt1G4412	1-401	g4432807	595	619	4.50E-59	88	8	(AC000106) Similar to Nicotiana protein kinase (gb D26601). [Arabidopsis thaliana]
1989	Bt1Gc6275	Bt1G4413	382-1	g4309763	475	494	6.20E-46	75	10	(AC006570) putative polyprotein [Arabidopsis thaliana]
1990	Bt1Gc6277	Bt1G4414	24-183	g1556378	110	154	3.60E-11	57	27	(AC006217) putative retrotransposon polyprotein [Arabidopsis thaliana]
1991	Bt1Gc6260	Bt1G4415	514-1	g3080399	372	376	7.50E-34	58	23	(Y08031) cheD [Bacillus cereus]
1992	Bt1Gc6287	Bt1G4417	128-1	g225559	116	138	1.60E-08	65	9	(AL022603) VP1 like protein [Arabidopsis thaliana]
1993	Bt1Gc6286	Bt1G4419	300-1	g1708181	202	207	2.30E-16	41	27	ORF IS231C [Bacillus thuringiensis]
1994	Bt1Gc6294	Bt1G4420	558-1	g228654	431	470	1.20E-44	46	28	PROBABLE OXYGEN-INDEPENDENT COPROPORPHYRINOGEN III OXIDASE (COPROPORPHYRINOGENASE) (COPROGEN OXIDASE) []
1995	Bt1Gc6299	Bt1G4421	426-178	g2392192	236	192	3.40E-15	54	36	thermostable pullulanase [Bacillus stearothermophilus]
1996	Bt1Gc6304	Bt1G4422	326-1	g730776	185	222	3.30E-17	43	14	Metallo-Beta-Lactamase From Bacillus Cereus 569H9 []
1997	Bt1Gc6323	Bt1G4424	1-161	g540914	233	111	1.30E-06	98	37	STAGE III SPORULATION PROTEIN E [Bacillus subtilis]
1997	Bt1Gc6323	Bt1G4425	167-457	g540915	480	487	1.90E-46	99	93	host-lethality protein kIbB - plasmid RK2 [Plasmid RK2]
1998	Bt1Gc6333	Bt1G4426	1-442	g118716	562	514	2.60E-49	67	39	trbD protein - plasmid RK2 [Plasmid RK2]
1999	Bt1Gc6340	Bt1G4427	54-230	g2984717	139	48	0.013	56	98	DNAJ PROTEIN [Bacillus subtilis]
2000	Bt1Gc6334	Bt1G4428	1-283	g3122814	88	124	3.50E-07	40	25	YisI homolog [Bacillus cereus]
2001	Bt1Gc6362	Bt1G4429	475-1	g1934609	312	328	2.10E-28	47	16	30S RIBOSOMAL PROTEIN S1 HOMOLOG [Bacillus cereus]
2002	Bt1Gc6373	Bt1G4430	1-274	g3288585	168	152	6.70E-10	45	16	formate dehydrogenase chain A [Bacillus subtilis]
2003	Bt1Gc6380	Bt1G4431	1-207	g3688229	76	75	0.14	36	10	preprotein translocase [Phormidium lamosum]
2004	Bt1Gc6386	Bt1G4432	423-1	g1945657	221	260	2.10E-22	40	54	(AJ011676) DNA ligase [Bacillus stearothermophilus]
2005	Bt1Gc6389	Bt1G4433	1-163	g1770035	109	80	0.034	45	9	hypothetical protein [Bacillus subtilis]
2006	Bt1Gc6401	Bt1G4434	1-214	g225559	142	179	5.80E-13	54	15	hypothetical protein [Bacillus subtilis]
2007	Bt1Gc6397	Bt1G4435	308-1	g2415397	158	139	1.90E-09	38	44	ORF IS231C [Bacillus thuringiensis]
2008	Bt1Gc6405	Bt1G4436	54-350	g1731036	117	144	3.10E-09	36	22	purine nucleoside phosphorylase [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2009	Bt1Gc6415	Bt1G4437	463-1	g2126871	338	351	4.90E-32	45	66	purine nucleoside phosphorylase (EC 2.7.-.-) II - Bacillus stearothermophilus [Bacillus stearothermophilus]
2010	Bt1Gc6419	Bt1G4438	362-37	g1075873	109	123	7.00E-08	30	59	nitrate reductase delta chain - Bacillus subtilis []
2011	Bt1Gc6420	Bt1G4439	471-310	g225268	121	102	1.50E-05	65	29	NADH dehydrogenase 2-like ORF 180 [Nicotiana tabacum]
2012	Bt1Gc6421	Bt1G4440	86-532	g2586081	223	271	1.10E-22	44	26	(U72725) receptor kinase-like protein [Oryza longistaminata]
2013	Bt1Gc6428	Bt1G4441	219-92	g1334366	167	184	2.40E-14	86	98	(X02441) gtg start [Nicotiana tabacum]
2014	Bt1Gc6433	Bt1G4442	1-201	g4263543	305	310	1.60E-26	88	7	(AC006250) putative Athila retroelement ORF1 protein [Arabidopsis thaliana]
2015	Bt1Gc6437	Bt1G4443	1-433	g538921	148	193	6.80E-15	32	40	aspartate-semialdehyde dehydrogenase (EC 1.2.1.11) - Bacillus subtilis []
2016	Bt1Gc6443	Bt1G4444	74-436	g2642163	548	540	1.50E-51	90	15	(AC003000) hypothetical protein [Arabidopsis thaliana]
2017	Bt1Gc6445	Bt1G4445	1-463	g4249386	92	128	4.30E-07	83	3	(AC005966) Strong similarity to gb AF061286 gamma-adaptin 1 from Arabidopsis thaliana. EST gb H37393 comes from this gene. [Arabidopsis thaliana]
2018	Bt1Gc6447	Bt1G4446	1-411	g1169301	445	421	1.90E-39	68	52	2,3-DIHYDRO-2,3-DIHYDROXYBENZOATE DEHYDROGENASE (COLD SHOCK PROTEIN CSI14) [Bacillus subtilis]
2019	Bt1Gc6449	Bt1G4447	556-1	g3063448	796	834	6.20E-83	89	15	(AC003981) F22O13.10 [Arabidopsis thaliana]
2020	Bt1Gc6452	Bt1G4448	138-422	g4539351	87	163	4.00E-11	54	11	(AL035539) putative protein [Arabidopsis thaliana]
2021	Bt1Gc6435	Bt1G4449	395-1	g2226127	141	221	2.90E-18	39	42	(Y14078) Hypothetical protein [Bacillus subtilis]
2022	Bt1Gc6456	Bt1G4450	1-347	g1673392	136	124	6.90E-07	35	20	(Z82044) unidentified transporter-ATP binding [Bacillus subtilis]
2023	Bt1Gc6460	Bt1G4451	162-412	g2462088	130	166	2.00E-12	45	45	(Y11170) BC541A protein [Bacillus cereus]
2024	Bt1Gc6457	Bt1G4452	1-218	g4773910	101	146	5.30E-10	63	18	(AF147259) No definition line found [Arabidopsis thaliana]
2025	Bt1Gc6465	Bt1G4453	47-267	g135188	121	154	1.30E-10	45	22	TRYPTOPHAN--TRNA SYNTHETASE (TRYPTOPHAN--TRNA LIGASE) (TRPRS) [Bacillus subtilis]
2026	Bt1Gc6466	Bt1G4454	1-375	g2127147	187	233	8.00E-19	41	25	nitrate reductase (EC 1.7.99.4) beta chain - Bacillus subtilis []
2027	Bt1Gc6478	Bt1G4455	1-354	g4033397	106	184	4.20E-13	35	14	DNA GYRASE SUBUNIT A [Bacillus sp.]
2028	Bt1Gc6487	Bt1G4456	1-266	g2245111	155	76	0.023	71	25	(Z97343) unnamed protein product [Arabidopsis thaliana]
2029	Bt1Gc6491	Bt1G4458	327-1	g1723607	111	127	3.40E-07	30	18	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN 2 IN GLVBC 3'REGION [Bacillus subtilis]
2030	Bt1Gc6494	Bt1G4459	271-1	g2635733	127	135	3.80E-09	40	89	(Z99120) yunC [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
2031	Bt1Gc6513	Bt1G4460	400-102	g4678942	230	293	6.80E-26	61	78	(AL049711) 5B protein like protein [Arabidopsis thaliana]
2032	Bt1Gc6527	Bt1G4461	352-1	g730965	214	211	4.00E-16	53	17	DNA TOPOISOMERASE I (OMEGA-PROTEIN) (RELAXING ENZYME) (UNTWISTING ENZYME) (SWIVELASE) [Bacillus subtilis]
2033	Bt1Gc6566	Bt1G4464	1-540	g134761	341	383	2.00E-35	48	52	STAGE II SPORULATION PROTEIN D [Bacillus subtilis]
2034	Bt1Gc6570	Bt1G4465	515-286	g3548811	294	320	9.40E-29	84	45	(AC005313) kinetochore (SKP1p)-like protein [Arabidopsis thaliana]
2035	Bt1Gc6572	Bt1G4466	1-400	g2618863	256	320	9.40E-29	45	51	(AF017113) YvpB [Bacillus subtilis]
2036	Bt1Gc6561	Bt1G4467	360-1	g4567262	353	336	1.90E-30	83	28	(AC006841) putative ubiquitin [Arabidopsis thaliana]
2037	Bt1Gc6560	Bt1G4468	1-277	g1731026	148	147	2.60E-09	35	14	HYPOTHETICAL 73.2 KD PROTEIN IN SODA-COMGA INTERGENIC REGION [Bacillus subtilis]
2038	Bt1Gc6573	Bt1G4469	1-430	g131137	240	321	7.20E-28	48	19	PHOTOSYSTEM I P700 CHLOROPHYLL A APOPROTEIN A1 [Pisum sativum]
2039	Bt1Gc6587	Bt1G4471	1-305	g2529467	256	326	2.20E-29	72	21	(AF006665) Yoks [Bacillus subtilis]
2040	Bt1Gc6590	Bt1G4472	1-408	g4678307	256	345	4.20E-31	87	14	(AL049655) vacuolar protein sorting-like protein [Arabidopsis thaliana]
2041	Bt1Gc6588	Bt1G4473	1-290	g136144	266	307	4.10E-27	66	20	TRANSPPOSASE FOR INSERTION SEQUENCE ELEMENT IS231C [Bacillus thuringiensis]
2042	Bt1Gc6599	Bt1G4474	584-1	g1703404	504	503	3.80E-48	49	69	SHIKIMATE 5-DEHYDROGENASE [Bacillus subtilis]
2043	Bt1Gc6617	Bt1G4476	317-1	g2314492	101	122	8.00E-07	30	23	(AE000634) fumarase (fumC) [Helicobacter pylori 26695]
2044	Bt1Gc6625	Bt1G4478	1-590	g4835229	470	427	4.30E-40	54	38	(AL049862) putative protein [Arabidopsis thaliana]
2045	Bt1Gc6628	Bt1G4480	292-1	g2293148	242	248	4.00E-21	52	36	(AF008220) dihydroxynaphthoate synthase [Bacillus subtilis]
2046	Bt1Gc6638	Bt1G4481	404-1	g1944414	242	257	6.00E-21	44	17	(D87026) glycogen phosphorylase [Bacillus stearothermophilus]
2047	Bt1Gc6641	Bt1G4482	293-69	g321919	83	106	4.50E-06	39	51	hypothetical 16.9K protein - Salmonella typhimurium plasmid NTP16 []
2048	Bt1Gc6639	Bt1G4483	1-294	g4530241	77	118	2.30E-06	32	20	(AF101234) D-alanine-D-alanyl carrier protein ligase DltA [Staphylococcus aureus]
2049	Bt1Gc6651	Bt1G4484	1-283	g2497382	119	154	2.40E-10	35	22	TRANSPPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
2050	Bt1Gc6652	Bt1G4485	478-1	g2293215	112	112	1.70E-05	29	23	(AF008220) YtpT [Bacillus subtilis]
2051	Bt1Gc6654	Bt1G4486	66-431	g2293212	254	299	1.60E-26	47	46	(AF008220) YtpQ [Bacillus subtilis]
2052	Bt1Gc6673	Bt1G4491	463-262	g2286111	130	196	1.30E-15	55	29	(U78891) MADS box protein [Oryza sativa]
2053	Bt1Gc6675	Bt1G4492	279-1	g4584140	130	145	7.10E-10	38	34	(AJ010139) fumarate hydratase [Bacillus cereus]



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2054	Bt1Gc6691	Bt1G4493	1-347	g2832244	367	183	2.00E-13	63	25	(AF031569) hypothetical protein [Zea mays]
2055	Bt1Gc6686	Bt1G4494	551-1	g479357	235	251	1.60E-20	65	14	hypothetical protein 612 - maize transposon MuA2 [Zea mays]
2056	Bt1Gc6700	Bt1G4495	120-1	g97193	96	105	9.60E-05	51	6	leukotoxin B - Pasteurella haemolytica []
2057	Bt1Gc6703	Bt1G4496	1-592	g2130141	254	184	4.20E-13	51	14	mudrA protein - maize transposon MuDR [Zea mays]
2058	Bt1Gc6709	Bt1G4499	1-275	g729583	119	121	6.10E-07	35	27	GLYCOGEN BIOSYNTHESIS PROTEIN GLGD [Bacillus subtilis]
2059	Bt1Gc6724	Bt1G4500	374-1	g2443256	205	221	1.90E-17	39	24	(D86417) YfmM [Bacillus subtilis]
2060	Bt1Gc6729	Bt1G4501	287-1	g2635764	103	145	3.30E-10	39	63	(Z99120) similar to NifU protein homolog [Bacillus subtilis]
2061	Bt1Gc6754	Bt1G4502	292-1	g482286	169	204	6.40E-15	41	7	DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain (version 2) - Bacillus subtilis []
2062	Bt1Gc6777	Bt1G4503	1-435	g2811062	126	178	7.60E-13	35	29	GLYCOGEN SYNTHASE (STARCH (BACTERIAL GLYCOGEN) SYNTHASE) [Bacillus stearothermophilus]
2063	Bt1Gc6786	Bt1G4504	1-169	g3328184	139	148	4.00E-10	59	19	(AF074376) beta lactamase precursor [Cloning vector pQE30NST]
2064	Bt1Gc6778	Bt1G4505	161-373	g4835325	102	109	9.40E-06	38	24	(AL049863) putative transcriptional regulator [Streptomyces coelicolor]
2065	Bt1Gc6798	Bt1G4506	406-1	g1731052	89	142	9.00E-09	32	17	PROBABLE 1-DEOXYXYLULOSE-5-PHOSPHATE SYNTHASE (DXP SYNTHASE) [Bacillus subtilis]
2066	Bt1Gc6802	Bt1G4507	214-1	g2293210	135	141	4.30E-09	43	20	(AF008220) YtoP [Bacillus subtilis]
2067	Bt1Gc6801	Bt1G4508	525-1	g3915461	272	308	8.50E-27	37	30	HYPOTHETICAL 66.6 KD PROTEIN IN PURD-SAPB INTERGENIC REGION [Bacillus subtilis]
2068	Bt1Gc6836	Bt1G4513	207-1	g3915204	95	111	3.20E-05	48	7	EXCINUCLEASE ABC SUBUNIT A [Bacillus subtilis]
2069	Bt1Gc6819	Bt1G4514	1-392	g1840129	148	180	8.10E-14	37	45	(U86347) beta-lactamase [Cloning vector pCALnFLAG]
2070	Bt1Gc6837	Bt1G4515	1-342	g3123297	130	118	4.90E-09	37	18	EXCINUCLEASE ABC SUBUNIT B (DNA PROTEIN) [Bacillus subtilis]
2071	Bt1Gc6842	Bt1G4516	1-630	g4914343	310	384	2.70E-34	43	19	(AC005489) F14N23.29 [Arabidopsis thaliana]
2072	Bt1Gc6849	Bt1G4518	508-403	g2529340	174	190	5.60E-15	97	14	(L81162) homologue; putative [Zea mays]
2073	Bt1Gc6871	Bt1G4521	129-405	g1002380	93	56	1	38	26	(U24189) RRM-type RNA binding protein [Caenorhabditis elegans]
2074	Bt1Gc6872	Bt1G4522	1-605	g3646371	163	208	6.90E-17	53	60	(AJ004916) chorismate mutase precursor [Prunus avium]
2075	Bt1Gc6867	Bt1G4524	1-629	g3522943	557	592	2.00E-56	57	16	(AC004411) putative p-glycoprotein [Arabidopsis thaliana]
2076	Bt1Gc6876	Bt1G4525	1-564	g4583548	655	645	3.40E-63	72	51	(AJ010820) FtsY homolog [Arabidopsis thaliana]
2077	Bt1Gc6881	Bt1G4526	1-641	g4415926	330	377	1.40E-33	43	20	(AC006282) unknown protein [Arabidopsis thaliana]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2078	Bt1Gc6890	Bt1G4537	529-367	g4263790	70	105	0.00031	46	3	(AC006068) putative ch-TOG protein [Arabidopsis thaliana]
2079	Bt1Gc6891	Bt1G4539	1-652	g4454463	113	127	4.70E-07	39	15	(AC006234) putative leucine rich protein kinase [Arabidopsis thaliana]
2080	Bt1Gc6902	Bt1G4542	1-310	g732345	149	178	1.00E-13	39	52	HYPOTHETICAL 22.6 KD PROTEIN IN EPR-GALK INTERGENIC REGION [Bacillus subtilis]
2081	Bt1Gc6909	Bt1G4544	1-409	g2959781	101	138	4.30E-08	29	14	(AJ223508) Zwillie protein [Arabidopsis thaliana]
2082	Bt1Gc6910	Bt1G4545	280-571	g3341694	148	194	2.10E-15	34	48	(AC003672) PREG-like protein [Arabidopsis thaliana]
2083	Bt1Gc6912	Bt1G4546	1-619	g4056432	719	759	2.80E-75	66	47	(AC005990) Similar to [Arabidopsis thaliana]
2084	Bt1Gc6922	Bt1G4548	286-1	g482286	141	211	1.10E-15	48	7	DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain (version 2) - Bacillus subtilis []
2085	Bt1Gc6927	Bt1G4549	142-630	g4115379	201	294	5.30E-26	40	53	(AC005967) putative carbonyl reductase [Arabidopsis thaliana]
2086	Bt1Gc6923	Bt1G4550	198-612	g4584429	289	295	4.20E-26	50	54	(AJ237751) aquaglyceroporin [Nicotiana tabacum]
2087	Bt1Gc6929	Bt1G4551	275-474	g4539386	57	97	0.00085	49	8	(AL035526) extensin-like protein [Arabidopsis thaliana]
2088	Bt1Gc6925	Bt1G4553	251-652	g2911073	184	254	8.00E-31	38	45	(AL021960) putative protein [Arabidopsis thaliana]
2089	Bt1Gc6956	Bt1G4558	1-308	g2072373	246	291	1.10E-25	59	30	(Y09252) phosphoribosylaminoimidazolecarboxamide formyltransferase [Bacillus cereus]
2090	Bt1Gc6951	Bt1G4561	377-1	g1648861	295	297	2.60E-26	48	29	(Z81356) UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Bacillus subtilis]
2091	Bt1Gc6957	Bt1G4562	1-625	g3954807	462	519	7.70E-50	52	40	(AJ011862) flavonoid 3',5'-hydroxylase [Catharanthus roseus]
2092	Bt1Gc6965	Bt1G4564	100-1	g225559	104	119	1.80E-06	71	7	ORF IS231C [Bacillus thuringiensis]
2093	Bt1Gc6980	Bt1G4565	1-188	g586883	106	132	7.80E-09	44	35	STAGE V SPORULATION PROTEIN T [Bacillus subtilis]
2094	Bt1Gc6974	Bt1G4567	19-362	g225559	372	388	5.80E-36	71	24	ORF IS231C [Bacillus thuringiensis]
2095	Bt1Gc6994	Bt1G4568	1-536	g2529340	738	804	4.80E-80	86	70	(L81162) homologue; putative [Zea mays]
2096	Bt1Gc7002	Bt1G4570	1-401	g134396	128	119	3.80E-06	30	16	PREPROTEIN TRANSLOCASE SECA SUBUNIT [Bacillus subtilis]
2097	Bt1Gc7015	Bt1G4573	1-522	g1352830	515	499	1.00E-47	64	31	VACUOLAR ATP SYNTHASE CATALYTIC SUBUNIT A (V-ATPASE 69 KD SUBUNIT) [Zea mays]
2098	Bt1Gc7023	Bt1G4574	267-1	g2635778	52	106	6.00E-05	41	11	(Z99120) similar to butyryl-CoA dehydrogenase [Bacillus subtilis]
2099	Bt1Gc7037	Bt1G4575	225-1	g3290176	105	118	1.50E-06	37	21	(AF067645) spore germination protein GerIB [Bacillus cereus]
2100	Bt1Gc7041	Bt1G4576	1-311	g1075694	147	190	4.80E-14	44	18	pheromone cAD1 binding protein precursor - Enterococcus

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2101	Bt1Gc7050	Bt1G4577	253-1	g126772	93	140	1.20E-08	41	15	faecalis plasmid pAD1 [Plasmid pAD1] MALATE SYNTHASE, GLYOXYSSOMAL [Neurospora crassa]
2102	Bt1Gc7056	Bt1G4578	1-340	g3929523	113	129	1.70E-07	35	21	(AF029224) NarH [Staphylococcus carnosus]
2103	Bt1Gc7064	Bt1G4579	1-296	g2506301	204	220	1.30E-17	55	23	ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPX [Bacillus subtilis]
2104	Bt1Gc7072	Bt1G4580	547-1	g732409	563	478	1.70E-45	59	55	HYPOTHETICAL 36.3 KD LIPOPROTEIN PRECURSOR (ORFK) [Bacillus subtilis]
2105	Bt1Gc7077	Bt1G4581	463-1	g1651978	116	159	7.00E-11	29	35	(D90901) hypothetical protein [Synechocystis sp.]
2106	Bt1Gc7103	Bt1G4583	146-484	g1653709	179	240	3.30E-19	48	16	(D90915) lipoprotein NlpD [Synechocystis sp.]
2107	Bt1Gc7110	Bt1G4584	1-171	g113368	101	101	9.00E-05	40	17	ALCOHOL DEHYDROGENASE I (ADH I) [Zymomonas mobilis]
2108	Bt1Gc7113	Bt1G4585	1-278	g141397	118	139	2.10E-09	45	38	HYPOTHETICAL OXIDOREDUCTASE IN RTP-PELB INTERGENIC REGION (ORF238) [Bacillus subtilis]
2109	Bt1Gc7121	Bt1G4586	1-313	g4467246	139	148	2.60E-09	35	14	(AL035569) putative phosphodiesterase [Streptomyces coelicolor]
2110	Bt1Gc7125	Bt1G4587	1-387	g1176995	232	262	1.30E-22	46	50	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN IN IDH 3'REGION [Bacillus subtilis]
2111	Bt1Gc7137	Bt1G4588	1-183	g2274867	141	179	2.00E-13	55	19	(AJ000005) glucose kinase [Bacillus megaterium]
2112	Bt1Gc7164	Bt1G4589	385-1	g1304006	124	162	4.20E-11	31	26	(D84432) Spo VAF [Bacillus subtilis]
2113	Bt1Gc7174	Bt1G4590	546-1	g556177	648	577	5.50E-56	69	44	(M222621) isocitrate lyase [Escherichia coli]
2114	Bt1Gc7175	Bt1G4591	349-1	g2984720	197	278	2.60E-24	56	55	(AF053927) probable spore germination protein C [Bacillus cereus]
2115	Bt1Gc7182	Bt1G4592	166-1	g3915700	102	111	7.30E-06	43	16	GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (GAPDH) [Bacillus stearothermophilus]
2116	Bt1Gc7187	Bt1G4593	1-416	g2635199	166	143	2.50E-10	42	30	(Z99117) similar to hypothetical proteins [Bacillus subtilis]
2117	Bt1Gc7224	Bt1G4594	1-125	g2493761	122	129	1.60E-08	60	61	MAJOR COLD SHOCK PROTEIN CSPA [Bacillus cereus]
2118	Bt1Gc7235	Bt1G4595	360-70	g120717	196	247	1.70E-20	52	21	GLUCOSE-6-PHOSPHATE ISOMERASE A (GPI A) (PHOSPHOGLUCOSE ISOMERASE A) [Bacillus stearothermophilus]
2119	Bt1Gc7236	Bt1G4596	1-430	g2632017	245	269	1.30E-22	40	27	(AJ002571) DppE [Bacillus subtilis]
2120	Bt1Gc7240	Bt1G4597	343-1	g2293126	101	134	1.50E-08	33	41	(U89962) beta-lactamase [Cloning vector pVP16]
2121	Bt1Gc7273	Bt1G4598	301-1	g135811	158	148	7.10E-10	47	28	THREONINE SYNTHASE [Bacillus sp.]
2122	Bt1Gc7277	Bt1G4599	1-420	g1934652	104	150	6.90E-10	32	31	(U93876) amino acid transporter [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2123	Bt1Gc7295	Bt1G4600	1-356	g2618835	203	166	4.10E-12	42	40	(AF017113) cell division protein [Bacillus subtilis]
2124	Bt1Gc7286	Bt1G4601	1-370	g4584142	319	349	7.90E-32	57	44	(AJ010139) DNA alkylation repair enzyme [Bacillus cereus]
2125	Bt1Gc7312	Bt1G4602	203-1	g114646	89	110	6.80E-06	38	24	ATP SYNTHASE GAMMA CHAIN PRECURSOR [thermophilic bacterium PS3]
2126	Bt1Gc7322	Bt1G4603	214-391	g114510	110	174	2.20E-12	64	11	ATP SYNTHASE ALPHA CHAIN [Bacillus megaterium]
2127	Bt1Gc7325	Bt1G4604	315-1	g114636	122	132	2.50E-08	37	37	ATP SYNTHASE GAMMA CHAIN [Bacillus megaterium]
2128	Bt1Gc7324	Bt1G4605	229-1	g114646	119	128	7.00E-08	40	27	ATP SYNTHASE GAMMA CHAIN PRECURSOR [thermophilic bacterium PS3]
2129	Bt1Gc7337	Bt1G4606	1-236	g4584142	169	145	8.90E-10	47	27	(AJ010139) DNA alkylation repair enzyme [Bacillus cereus]
2130	Bt1Gc7350	Bt1G4607	1-174	g2497382	130	155	1.90E-10	56	13	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
2131	Bt1Gc7345	Bt1G4608	1-430	g2149596	87	139	1.90E-08	47	13	(U89796) chitinase [Bacillus thuringiensis]
2132	Bt1Gc7347	Bt1G4609	1-375	g1175720	152	202	3.00E-16	37	42	HYPOTHETICAL OXIDOREDUCTASE IN PBPC-LRPC INTERGENIC REGION [Bacillus subtilis]
2133	Bt1Gc7352	Bt1G4610	438-1	g482286	263	285	1.50E-23	44	10	DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain (version 2) - Bacillus subtilis []
2134	Bt1Gc7357	Bt1G4611	1-489	g1352095	212	174	1.00E-12	38	46	COMG OPERON PROTEIN 1 [Bacillus subtilis]
2135	Bt1Gc7361	Bt1G4612	1-351	g2497382	174	229	1.50E-18	49	27	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
2136	Bt1Gc7375	Bt1G4613	345-1	g1346442	95	130	1.40E-07	29	22	2-ISOPROPYLMALATE SYNTHASE (ALPHA-ISOPROPYLMALATE SYNTHASE) (ALPHA-IPM SYNTHETASE) [Anabaena PCC7120]
2137	Bt1Gc7374	Bt1G4614	218-1	g2492562	140	116	3.90E-07	38	53	EXCINUCLEASE ABC SUBUNIT A [Streptococcus mutans]
2138	Bt1Gc7397	Bt1G4615	248-48	g3955202	115	69	0.081	39	39	(AF022796) MoaB [Staphylococcus carnosus]
2139	Bt1Gc7396	Bt1G4616	1-308	g482286	186	220	1.30E-16	46	7	DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain (version 2) - Bacillus subtilis []
2140	Bt1Gc7405	Bt1G4617	334-1	g2635763	227	265	2.00E-22	53	24	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
2141	Bt1Gc7400	Bt1G4618	530-1	g2415745	211	281	1.30E-24	40	47	(AB000617) YceH [Bacillus subtilis]
2142	Bt1Gc7426	Bt1G4619	403-1	g2492639	140	210	2.50E-16	38	28	3-ISOPROPYLMALATE DEHYDRATASE LARGE SUBUNIT (ISOPROPYLMALATE ISOMERASE) (ALPHA-IPM ISOMERASE) (IPMI) [Actinoplanes teichomyceticus]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2143	Bt1Gc7429	Bt1G4620	1-272	g72992	179	131	5.10E-08	49	26	recE protein - Bacillus subtilis []
2144	Bt1Gc7440	Bt1G4621	1-423	g1731004	262	297	2.60E-26	46	44	LYTB PROTEIN HOMOLOG [Bacillus subtilis]
2145	Bt1Gc7392	Bt1G4622	1-264	g2293164	144	191	1.90E-14	53	22	(AF008220) SAM synthase [Bacillus subtilis]
2146	Bt1Gc7470	Bt1G4623	143-1	g1934831	100	119	9.20E-07	52	14	(Z93939) unknown [Bacillus subtilis]
2147	Bt1Gc7473	Bt1G4624	1-268	g135106	161	140	9.80E-09	43	19	GLUTAMYL-TRNA SYNTHETASE (GLUTAMATE--TRNA LIGASE) (GLURS) [Bacillus subtilis]
2148	Bt1Gc7488	Bt1G4625	1-547	g549706	120	205	1.50E-16	35	45	KTI12 PROTEIN [Saccharomyces cerevisiae]
2149	Bt1Gc7498	Bt1G4626	1-533	g1351995	504	525	3.70E-49	58	12	PROBABLE CALCIUM-TRANSPORTING ATPASE [Schizosaccharomyces pombe]
2150	Bt1Gc7495	Bt1G4627	1-275	g1945118	373	372	2.90E-34	82	84	(D88802) groES [Bacillus subtilis]
2150	Bt1Gc7495	Bt1G4628	317-504	g2119971	247	276	2.20E-23	88	12	molecular chaperone 60 GroEL - Bacillus sp []
2151	Bt1Gc7500	Bt1G4629	1-540	g3925779	238	181	5.00E-13	29	31	(AL034353) putative major facilitator family multi-drug resistance protein [Schizosaccharomyces pombe]
2152	Bt1Gc7508	Bt1G4630	1-223	g2501412	124	155	1.70E-10	47	18	TRYPTOPHAN SYNTHASE BETA CHAIN [Methanococcus jannaschii]
2153	Bt1Gc7507	Bt1G4631	168-1	g2497382	130	187	6.20E-14	61	13	TRANSPPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
2154	Bt1Gc7583	Bt1G4632	1-311	g2815006	192	230	5.10E-18	41	13	(AJ001087) pullulanase [Thermotoga maritima]
2155	Bt1Gc7592	Bt1G4633	265-1	g728788	177	208	6.00E-16	50	15	ACETYL-COENZYME A SYNTHETASE (ACETATE--COA LIGASE) (ACYL-ACTIVATING ENZYME) (ACETYL-COA SYNTHASE) [Bacillus subtilis]
2156	Bt1Gc7590	Bt1G4634	421-1	g3914289	216	197	1.20E-14	44	21	TOPOISOMERASE IV SUBUNIT B [Bacillus subtilis]
2157	Bt1Gc7597	Bt1G4635	62-274	g2636123	128	188	9.10E-15	50	77	(Z99122) ywsA [Bacillus subtilis]
2158	Bt1Gc7616	Bt1G4636	363-1	g3122850	184	204	1.10E-15	40	26	PREPROTEIN TRANSLOCASE SECA SUBUNIT [Bacillus firmus]
2159	Bt1Gc7619	Bt1G4637	255-1	g585047	106	148	8.40E-10	39	23	ALANINE DEHYDROGENASE (STAGE V SPORULATION PROTEIN N) [Bacillus subtilis]
2160	Bt1Gc7625	Bt1G4638	254-1	g2635763	156	177	9.00E-13	40	18	(Z99120) similar to hypothetical proteins [Bacillus subtilis]
2161	Bt1Gc7651	Bt1G4639	416-1	g2633126	168	224	6.80E-18	36	30	(Z99108) similar to RNA methyltransferase [Bacillus subtilis]
2162	Bt1Gc7653	Bt1G4640	1-186	g3183527	104	120	1.50E-07	40	39	TRANSCRIPTION ELONGATION FACTOR GREY (TRANSCRIPT CLEAVAGE FACTOR GREY) (GENERAL STRESS PROTEIN 20M) (GSP20M) [Bacillus subtilis]
2163	Bt1Gc7654	Bt1G4641	492-1	g584920	723	493	4.40E-47	89	30	60 KD CHAPERONIN (PROTEIN CPN60) (GROEL)

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2164	Bt1Gc7674	Bt1G4643	1-296	g2634069	131	154	6.00E-11	37	37	PROTEIN) [Bacillus stearothermophilus]
2165	Bt1Gc7677	Bt1G4644	446-2	g2633912	221	266	4.90E-23	40	99	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
2166	Bt1Gc7693	Bt1G4645	413-52	g2462097	95	174	2.80E-13	36	47	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
2167	Bt1Gc7699	Bt1G4646	101-479	g460277	439	282	1.00E-24	81	64	(Y10927) glutamate racemase [Bacillus cereus]
2168	Bt1Gc7707	Bt1G4647	184-1	g2500611	116	140	2.90E-08	44	6	(L29404) ORF; putative [Escherichia coli]
2169	Bt1Gc7708	Bt1G4648	1-383	g400783	122	165	2.50E-12	33	52	DNA-DIRECTED RNA POLYMERASE BETA' CHAIN (RNA TRANSCRIPTASE BETA' CHAIN) (RNA POLYMERASE BETA' SUBUNIT) []
2170	Bt1Gc7713	Bt1G4649	147-1	g3122885	99	121	1.50E-06	52	8	ALKALINE PHOSPHATASE SYNTHESIS TRANSCRIPTIONAL REGULATORY PROTEIN PHOP [Bacillus subtilis]
2171	Bt1Gc7716	Bt1G4650	1-194	g3402816	82	132	2.20E-08	40	24	ASPARTYL-TRNA SYNTHETASE (ASPARTATE--TRNA LIGASE) (ASPRS) [Bacillus subtilis]
2172	Bt1Gc7715	Bt1G4651	257-376	g4512401	196	196	1.70E-14	95	6	(AJ007829) NptI [Cloning vector pGreen]
2173	Bt1Gc7717	Bt1G4652	471-1	g125935	168	205	1.30E-15	29	27	(AB017508) fus homologue (identity of 87% to B. subtilis ) [Bacillus halodurans]
2174	Bt1Gc7723	Bt1G4653	542-271	g3915194	92	136	2.90E-09	33	71	LACTOSE PERMEASE [Kluyveromyces lactis]
2175	Bt1Gc7719	Bt1G4654	259-1	g2851553	115	128	2.40E-07	36	15	HYPOTHETICAL TRANSTHYRETIN-LIKE PROTEIN R09H10.3 IN CHROMOSOME IV [Caenorhabditis elegans]
2176	Bt1Gc7743	Bt1G4655	188-1	g2239289	127	139	1.10E-08	42	14	PROBABLE PHOSPHOMANNOMUTASE (PMM) [Bacillus subtilis]
2177	Bt1Gc7760	Bt1G4657	1-561	g128340	94	164	6.40E-11	35	8	(U51115) unknown protein [Bacillus subtilis]
2178	Bt1Gc7773	Bt1G4658	1-190	g1805377	111	119	8.90E-07	44	20	NITROGEN ASSIMILATION TRANSCRIPTION FACTOR NIRA [Emericella nidulans]
2179	Bt1Gc7777	Bt1G4659	1-348	g2632985	98	107	9.60E-05	30	11	(D50453) L-lactate dehydrogenase [Bacillus subtilis]
2180	Bt1Gc7778	Bt1G4660	1-131	g3021327	116	138	2.10E-08	68	8	(Z99107) similar to acriflavin resistance protein [Bacillus subtilis]
2181	Bt1Gc7789	Bt1G4661	1-327	g1064785	101	108	7.50E-06	28	45	(AJ005075) enzyme I [Bacillus megaterium]
2182	Bt1Gc7791	Bt1G4662	1-384	g3183185	400	427	4.30E-40	63	21	(D78193) yydK [Bacillus subtilis]
2183	Bt1Gc7804	Bt1G4663	381-82	g886038	80	143	6.40E-10	40	39	GTP-BINDING PROTEIN TYPA/BIPA HOMOLOG [Bacillus subtilis]
2184	Bt1Gc7809	Bt1G4664	198-1	g4512355	102	116	3.90E-07	42	49	(U24659) ladR1 [Streptomyces venezuelae]
										(AB011836) similar to B.subtilis ywgb gene(27%-identity) [Bacillus halodurans]

### NCBI gi description

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2185	Bt1Gc7767	Bt1G4665	285-1	g586022	185	208	7.50E-16	46	15	STAGE V SPOULATION PROTEIN D (SPOULATION SPECIFIC PENICILLIN-BINDING PROTEIN) [Bacillus subtilis]
2186	Bt1Gc7814	Bt1G4666	1-314	g132246	131	159	5.00E-11	35	28	RECF PROTEIN [Bacillus subtilis]
2187	Bt1Gc7793	Bt1G4667	428-79	g2633216	228	170	7.40E-13	42	73	(Z99108) rRNA methylase homolog [Bacillus subtilis]
2188	Bt1Gc7823	Bt1G4668	834-1	g585648	390	457	8.20E-43	42	30	PENICILLIN-BINDING PROTEIN 2B (PBP-2B) [Bacillus subtilis]
2189	Bt1Gc7828	Bt1G4669	384-1	g225559	74	98	0.00032	34	18	ORF IS231C [Bacillus thuringiensis]
2190	Bt1Gc7836	Bt1G4670	163-1	g2632519	100	124	2.50E-07	44	17	(Z99105) similar to hypothetical proteins [Bacillus subtilis]
2191	Bt1Gc7852	Bt1G4673	1-269	g1709188	151	163	4.90E-11	40	14	DNA MISMATCH REPAIR PROTEIN MUTL [Bacillus subtilis]
2192	Bt1Gc7853	Bt1G4674	1-164	g124464	137	159	1.50E-10	62	8	IMMUNE INHIBITOR A PRECURSOR [Bacillus thuringiensis]
2193	Bt1Gc7856	Bt1G4675	1-387	g2582651	188	241	2.20E-20	45	34	(AJ004803) putative undecaprenyl-phosphate N-acetylglucosaminyltransferase [Bacillus subtilis]
2194	Bt1Gc7858	Bt1G4676	1-254	g732387	353	360	5.40E-33	78	56	HYPOTHETICAL LACA/RPIB FAMILY PROTEIN IN SPOIR-GLYC INTERGENIC REGION [Bacillus subtilis]
2195	Bt1Gc7854	Bt1G4677	120-258	g115626	81	128	5.60E-07	49	5	CARBAMOYL-PHOSPHATE SYNTHASE, PYRIMIDINE-SPECIFIC, LARGE CHAIN (CARBAMOYL-PHOSPHATE SYNTHETASE AMMONIA CHAIN) [Bacillus subtilis]
2196	Bt1Gc7868	Bt1G4678	372-547	g1708267	170	186	6.50E-14	64	15	FLAVOHEMOPROTEIN (HAEMOGLOBIN-LIKE PROTEIN) (FLAVOHEMOGLOBIN) [Bacillus subtilis]
2197	Bt1Gc7863	Bt1G4679	333-1	g2634201	88	116	1.60E-06	27	37	(Z99113) similar to UTP-glucose-1-phosphate uridylyltransferase [Bacillus subtilis]
2198	Bt1Gc7879	Bt1G4680	1-210	g132760	245	252	1.50E-21	70	58	50S RIBOSOMAL PROTEIN L20 [Bacillus stearothermophilus]
2199	Bt1Gc7916	Bt1G4681	474-1	g2633814	554	582	2.70E-59	73	29	(Z99111) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis]
2200	Bt1Gc7934	Bt1G4682	1-361	g1723607	173	192	3.50E-14	37	20	HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN 2 IN GLVBC 3'REGION [Bacillus subtilis]
2201	Bt1Gc7951	Bt1G4683	123-1	g586900	115	134	9.00E-08	62	5	NEGATIVE REGULATOR OF GENETIC COMPETENCE MECB [Bacillus subtilis]
2202	Bt1Gc7956	Bt1G4685	268-1	g1731040	110	132	8.90E-08	39	16	HYPOTHETICAL HELICASE IN SINI-GCVT INTERGENIC REGION [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2203	Bt1Gc7954	Bt1G4686	61-312	g321919	103	102	1.20E-05	35	56	hypothetical 16.9K protein - Salmonella typhimurium plasmid NTP16 []
2204	Bt1Gc7965	Bt1G4687	78-239	g2632453	179	188	9.10E-15	65	31	(Z99104) ybcF [Bacillus subtilis]
2205	Bt1Gc7990	Bt1G4688	1-212	g3183561	174	219	4.70E-18	63	26	VEGETATIVE PROTEIN 296 (VEG296) [Bacillus subtilis]
2206	Bt1Gc8000	Bt1G4689	62-332	g2145613	178	197	1.00E-15	47	99	5'-phosphoribosyl-glycinamide synthetase (EC 6.3.4.13) (fragment) - Listeria monocytogenes [Listeria monocytogenes]
2207	Bt1Gc8012	Bt1G4690	1-383	g755588	247	284	1.00E-23	47	13	(U23149) DNA polymerase [Bacillus stearothermophilus]
2208	Bt1Gc8013	Bt1G4691	67-468	g3183562	148	195	5.90E-15	43	30	SUCCINYL-COA SYNTHETASE BETA CHAIN (SCS-ALPHA) (VEGETATIVE PROTEIN 63) (VEG63) [Bacillus subtilis]
2209	Bt1Gc8024	Bt1G4692	1-388	g1770066	138	160	9.00E-11	33	22	(Z75208) acetolactate synthase large subunit [Bacillus subtilis]
2210	Bt1Gc8044	Bt1G4694	443-1	g2633912	203	230	3.20E-19	39	97	(Z99112) similar to hypothetical proteins [Bacillus subtilis]
2211	Bt1Gc8055	Bt1G4695	228-1	g4416482	101	115	2.40E-06	38	24	(AF125999) daunorubicin resistance protein A [Mycobacterium avium]
2212	Bt1Gc8048	Bt1G4696	1-286	g133475	272	298	2.00E-26	63	36	RNA POLYMERASE SIGMA-G FACTOR (STAGE III SPORULATION PROTEIN G) [Bacillus subtilis]
2213	Bt1Gc8064	Bt1G4697	335-1	g586022	176	182	4.70E-13	37	17	STAGE V SPORULATION PROTEIN D (SPORULATION SPECIFIC PENICILLIN-BINDING PROTEIN) [Bacillus subtilis]
2214	Bt1Gc8067	Bt1G4698	1-168	g2632231	109	130	1.30E-08	45	70	(AJ222587) YkuJ protein [Bacillus subtilis]
2215	Bt1Gc8059	Bt1G4699	19-267	g2633523	166	176	1.70E-13	51	32	(Z99110) similar to thiamin biosynthesis [Bacillus subtilis]
2216	Bt1Gc8071	Bt1G4700	289-1	g417115	187	200	2.90E-15	42	20	PROTOPORPHYRINOGEN OXIDASE (PPO) [Bacillus subtilis]
2217	Bt1Gc8074	Bt1G4701	164-1	g1929340	104	126	4.10E-07	45	9	(Z93767) alpha-acetolactate synthase protein, AlsS [Bacillus subtilis]
2218	Bt1Gc8076	Bt1G4702	200-21	g2127299	144	167	1.50E-12	55	42	hypothetical protein B - Clostridium acetobutylicum [Clostridium acetobutylicum]
2219	Bt1Gc8083	Bt1G4703	253-1	g2851553	78	112	1.30E-05	36	15	PROBABLE PHOSPHOMANNOMUTASE (PMM) [Bacillus subtilis]
2220	Bt1Gc8099	Bt1G4704	310-1	g226781	68	125	1.60E-06	35	6	RNA polymerase beta' [Pseudomonas putida]
2221	Bt1Gc8098	Bt1G4705	1-179	g3152725	91	126	2.70E-07	46	12	(AF065394) enolase [Staphylococcus aureus]
2222	Bt1Gc8102	Bt1G4706	256-1	g730399	159	169	1.10E-11	47	13	PKA PROTEIN [Bacillus subtilis]
2223	Bt1Gc8106	Bt1G4707	1-383	g732327	355	384	1.50E-35	60	41	PROBABLE 1,4-DIHYDROXY-2-NAPHTHOATE



Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aaf_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2224	Bt1Gc8109	Bt1G4708	283-1	g1351856	66	114	1.40E-05	35	10	OCTAPRENYLTRANSFERASE (DHNA- OCTAPRENYLTRANSFERASE) [Bacillus subtilis]
2225	Bt1Gc8124	Bt1G4709	377-1	g399891	82	152	5.90E-11	31	60	ACONITATE HYDRATASE, CYTOPLASMIC (CITRATE HYDRO-LYASE) (ACONITASE) [Cucurbita sp.]
2226	Bt1Gc8120	Bt1G4710	449-1	g2829488	253	300	1.20E-26	47	35	ATP PHOSPHORIBOSYLTRANSFERASE [Lactococcus lactis]
2227	Bt1Gc8130	Bt1G4711	437-1	g1710618	94	84	0.0041	40	46	GLUTAMATE-1-SEMIALDEHYDE 2,1-AMINOMUTASE 2 (GSA) (GLUTAMATE-1-SEMIALDEHYDE AMINOTRANSFERASE) (GSA-AT) [Bacillus subtilis]
2228	Bt1Gc8133	Bt1G4712	223-80	g4512408	241	241	2.20E-20	92	51	PROBABLE RIBONUCLEASE PH (RNASE PH) (TRNA NUCLEOTIDYLTRANSFERASE) [Mycobacterium tuberculosis]
2229	Bt1Gc8140	Bt1G4713	1-287	g1351054	85	106	9.20E-05	27	11	(AB017508) rpsS homologue (identity of 87% to B. subtilis ) [Bacillus halodurans]
2230	Bt1Gc8172	Bt1G4716	372-432	g1255196	93	109	2.80E-05	90	3	PREPROTEIN TRANSLOCASE SECA SUBUNIT [Listeria monocytogenes]
2231	Bt1Gc8177	Bt1G4717	293-1	g1945649	118	145	3.30E-10	32	62	(U50744) BSMA [Bacillus stearothermophilus]
2232	Bt1Gc8178	Bt1G4718	1-240	g729328	259	269	2.40E-23	64	31	(Z94043) hypothetical protein [Bacillus subtilis]
2233	Bt1Gc8186	Bt1G4719	360-1	g2500058	147	137	3.80E-08	34	16	GLUCOSE 1-DEHYDROGENASE [Bacillus megaterium]
2234	Bt1Gc8199	Bt1G4720	1-336	g2127280	77	123	2.40E-08	27	30	FORMATE ACETYLTRANSFERASE (PYRUVATE FORMATE-LYASE) [Clostridium pasteurianum]
2235	Bt1Gc8222	Bt1G4721	389-1	g135196	192	224	4.80E-18	39	31	histidine protein kinase (Tn5401) - Bacillus thuringiensis [Bacillus thuringiensis]
2236	Bt1Gc8230	Bt1G4723	1-550	g1075694	293	349	1.60E-31	38	34	TYROSYL-TRNA SYNTHETASE (TYROSINE--TRNA LIGASE) (TYRRS) [Bacillus caldotenax]
2237	Bt1Gc8231	Bt1G4724	206-1	g4582216	218	227	7.70E-18	67	10	pheromone cAD1 binding protein precursor - Enterococcus faecalis plasmid pAD1 [Plasmid pAD1]
2238	Bt1Gc8242	Bt1G4725	260-1	g464794	104	204	1.20E-15	48	18	(AJ237696) elongation factor G (EF-G) [Staphylococcus aureus]
2239	Bt1Gc8243	Bt1G4726	280-1	g2116761	196	103	7.50E-05	37	23	STAGE IV SPORULATION PROTEIN A [Bacillus subtilis]
2240	Bt1Gc8245	Bt1G4727	450-1	g2960100	83	124	9.10E-08	22	67	(D86418) YfnC [Bacillus subtilis]
2241	Bt1Gc8244	Bt1G4728	335-1	g2635766	135	183	1.70E-13	41	26	(AL022121) hypothetical protein Rv3676 [Mycobacterium tuberculosis] similar to hypothetical proteins [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aaf <sub>nap</sub> Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
2242	Bt1Gc8252	Bt1G4729	423-45	g124464	475	518	9.80E-50	78	18	IMMUNE INHIBITOR A PRECURSOR [Bacillus thuringiensis]
2243	Bt1Gc8258	Bt1G4730	322-1	g1894751	199	216	9.80E-18	40	37	(Z92952) ywqM [Bacillus subtilis]
2244	Bt1Gc8263	Bt1G4731	1-279	g2635778	168	203	2.20E-15	53	15	(Z99120) similar to butyryl-CoA dehydrogenase [Bacillus subtilis]
2245	Bt1Gc8282	Bt1G4732	124-256	g2619056	161	93	0.001	78	10	(AF027868) putative transporter [Bacillus subtilis]
2246	Bt1Gc8280	Bt1G4733	355-1	g482286	135	160	3.10E-10	40	8	DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain (version 2) - Bacillus subtilis []
2247	Bt1Gc8290	Bt1G4734	1-404	g2635780	165	142	1.30E-08	39	16	(Z99120) similar to 3-hydroxyacyl-CoA dehydrogenase [Bacillus subtilis]
2248	Bt1Gc8286	Bt1G4735	154-22	g1196998	108	118	1.70E-06	61	11	(J01829) unknown protein [Transposon Tn10]
2249	Bt1Gc8308	Bt1G4736	388-155	g1575605	111	142	7.30E-09	40	14	(U68411) 4-methyl-5-nitrocatechol oxygenase [Burkholderia sp.]
2250	Bt1Gc8319	Bt1G4737	1-463	g3024353	168	185	2.70E-13	37	22	ATP-DEPENDENT HELICASE PCRA []
2251	Bt1Gc8318	Bt1G4738	386-1	g3688229	177	207	1.00E-15	40	19	(AJ011676) DNA ligase [Bacillus stearothermophilus]
2252	Bt1Gc8335	Bt1G4739	382-1	g1168649	154	182	3.90E-14	38	49	MENAJUINOL-CYTOCHROME C REDUCTASE CYTOCHROME B/C SUBUNIT [Bacillus subtilis]
2253	Bt1Gc8347	Bt1G4740	1-293	g2116760	112	133	9.90E-09	36	41	(D86418) YfjB [Bacillus subtilis]
2254	Bt1Gc8353	Bt1G4742	1-205	g3123227	180	118	2.50E-06	55	13	GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE) [Bacillus subtilis]
2255	Bt1Gc8358	Bt1G4743	436-193	g2293166	209	151	5.90E-10	53	17	(AF008220) amino acid transporter [Bacillus subtilis]
2256	Bt1Gc8368	Bt1G4744	366-1	g1346148	96	130	2.50E-08	31	52	GLUTAMATE TRANSPORT ATP-BINDING PROTEIN GLUA [Corynebacterium glutamicum]
2257	Bt1Gc8372	Bt1G4745	1-485	g586900	119	194	3.50E-14	39	20	NEGATIVE REGULATOR OF GENETIC COMPETENCE MECA [Bacillus subtilis]
2258	Bt1Gc8390	Bt1G4746	1-306	g3025119	163	181	1.50E-13	43	29	HYPOTHETICAL 36.8 KD PROTEIN IN PHOB-GROES INTERGENIC REGION [Bacillus subtilis]
2259	Bt1Gc8392	Bt1G4747	412-1	g1825778	53	134	3.80E-08	30	31	(U88315) similar to acetyltransferases [Caenorhabditis elegans]
2260	Bt1Gc8397	Bt1G4748	1-475	g729157	132	171	4.00E-12	34	34	COMF OPERON PROTEIN 1 [Bacillus subtilis]
2261	Bt1Gc8391	Bt1G4749	176-1	g2506131	115	165	5.10E-11	56	7	ACONITATE HYDRATASE (CITRATE HYDRO-LYASE) (ACONITASE) [Bacillus subtilis]
2262	Bt1Gc8415	Bt1G4750	1-340	g730600	288	200	2.90E-15	56	24	AMINO-ACID PERMEASE ROCC [Bacillus subtilis]
2263	Bt1Gc8420	Bt1G4751	160-1	g1765902	91	112	1.80E-06	46	25	(X99545) uracil phosphoribosyltransferase [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2264	Bt1Gc8417	Bt1G4752	1-405	g1177011	440	415	8.00E-39	65	31	caldolyticus] HYPOTHETICAL 47.3 KD PROTEIN IN WAPA-LICT INTERGENIC REGION [Bacillus subtilis]
2265	Bt1Gc8427	Bt1G4753	555-1	g2209087	122	176	2.20E-12	33	23	(AF000309) putative serine/threonine kinase [Colletotrichum lindemuthianum]
2266	Bt1Gc8433	Bt1G4754	514-1	g4587971	820	822	3.20E-81	97	13	(AF082072) ABC transporter protein AtrC [Emericella nidulans]
2267	Bt1Gc8431	Bt1G4755	350-109	g2506915	121	59	0.88	41	32	OUTER DENSE FIBER PROTEIN (RT7 PROTEIN) (RTS 5/1) [Rattus norvegicus]
2268	Bt1Gc8435	Bt1G4756	317-1	g3290177	204	175	7.90E-13	48	29	(AF067645) spore germination protein GerIC [Bacillus cereus]
2269	Bt1Gc8438	Bt1G4757	361-1	g710552	63	124	3.00E-06	32	5	(L40632) ankyrin 3 [Mus musculus]
2270	Bt1Gc8444	Bt1G4758	1-387	g135197	175	103	7.70E-05	41	31	TYROSYL-TRNA SYNTHETASE (TYROSINE--TRNA LIGASE) (TYRRS) []
2271	Bt1Gc8439	Bt1G4760	571-1	g3192023	177	241	1.10E-19	37	27	(AL023794) putative proline oxidase precursor [Schizosaccharomyces pombe]
2272	Bt1Gc8457	Bt1G4762	377-1	g2293156	88	106	5.70E-06	29	67	(AF008220) YtiB [Bacillus subtilis]
2273	Bt1Gc8470	Bt1G4764	393-1	g728788	217	258	2.30E-21	40	23	ACETYL-COENZYME A SYNTHETASE (ACETATE--COA LIGASE) (ACYL-ACTIVATING ENZYME) (ACETYL-COA SYNTHASE) [Bacillus subtilis]
2274	Bt1Gc8468	Bt1G4765	1-434	g3879121	122	183	2.20E-12	34	6	(Z70310) predicted using Genefinder; Similarity to Mouse ankyrin (PIR Acc. No. S37771); cDNA EST EMBL:T01923 comes from this gene; cDNA EST EMBL:D32335 comes from this gene; cDNA EST EMBL:D32723 comes from this gene; cDNA EST E... []
2275	Bt1Gc8475	Bt1G4766	1-317	g1730878	419	427	8.50E-40	80	16	HYPOTHETICAL 69.1 KD PROTEIN (ORF4) [Corynebacterium glutamicum]
2276	Bt1Gc8476	Bt1G4767	458-1	g2342601	252	296	4.70E-24	42	3	(X89442) peptide synthetase [Metarhizium anisopliae]
2277	Bt1Gc8480	Bt1G4768	1-436	g2415745	218	269	2.40E-23	42	40	(AB000617) YceH [Bacillus subtilis]
2278	Bt1Gc8491	Bt1G4769	1-493	g96497	821	759	2.80E-75	98	93	transcriptional repressor korFII - plasmid RK2 []
2279	Bt1Gc8497	Bt1G4770	166-1	g2337805	97	116	1.10E-06	46	22	(Y13937) putative Ptc1 protein [Bacillus subtilis]
2280	Bt1Gc8484	Bt1G4771	1-434	g586058	139	257	2.00E-21	43	28	LYSYL-TRNA SYNTHETASE (LYSINE--TRNA LIGASE) (LYSRS) [Bacillus subtilis]
2281	Bt1Gc8509	Bt1G4772	425-1	g4126672	168	136	1.20E-08	31	43	(AB016431) czcD [Staphylococcus aureus]
2282	Bt1Gc8513	Bt1G4773	390-1	g2266425	198	226	3.50E-18	40	29	(Y13917) yngH [Bacillus subtilis]

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aat_nap Score	BlastP Score	BlastP-Prob	% Ident	% Cvr	NCBI gi description
2283	Bt1Gc8540	Bt1G4774	1-207	g2688299	84	119	1.80E-06	39	14	(AE001145) prolyl-tRNA synthetase (proS) [Borrelia burgdorferi]
2284	Bt1Gc8531	Bt1G4775	417-1	g1881236	172	199	8.50E-15	42	19	(AB001488) PROBABLE DNA TOPOISOMERASE III [Bacillus subtilis]
2285	Bt1Gc8541	Bt1G4776	576-1	g1351714	112	181	3.80E-13	22	35	PUTATIVE TRANSPORTER C11D3.18C [Schizosaccharomyces pombe]
2286	Bt1Gc8552	Bt1G4778	1-206	g4584073	95	118	2.40E-07	40	41	(AJ236899) hypothetical protein [Streptococcus gordonii]
2287	Bt1Gc8567	Bt1G4779	561-1	g3135999	98	187	6.60E-14	29	34	(AL023589) hypothetical protein [Schizosaccharomyces pombe]
2288	Bt1Gc8568	Bt1G4780	1-500	g134506	149	264	4.90E-22	42	23	REGULATORY PROTEIN SIR2 (SILENT INFORMATION REGULATOR 2) [Saccharomyces cerevisiae]
2289	Bt1Gc8574	Bt1G4781	1-482	g3123227	487	492	5.60E-47	60	31	GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE) [Bacillus subtilis]
2290	Bt1Gc8575	Bt1G4782	1-105	g464933	168	168	1.20E-12	97	23	TRAM PROTEIN [Escherichia coli]
2291	Bt1Gc8576	Bt1G4783	280-1	g120474	133	109	8.10E-06	36	34	FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (FAPY-DNA GLYCOSYLASE) [Bacillus firmus]
2292	Bt1Gc8580	Bt1G4784	408-1	g135103	70	138	8.30E-08	26	8	MULTIFUNCTIONAL AMINOACYL-TRNA SYNTHETASE [Drosophila melanogaster]
2293	Bt1Gc8590	Bt1G4785	304-1	g80261	111	149	4.60E-09	34	7	DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain (version 1) - Bacillus subtilis []
2294	Bt1Gc8589	Bt1G4786	1-181	g2293212	187	188	9.10E-15	61	22	(AF008220) YtpQ [Bacillus subtilis]
2295	Bt1Gc8594	Bt1G4787	297-1	g2293328	218	238	4.60E-20	44	35	(AF008220) putative morphine dehydrogenase [Bacillus subtilis]
2296	Bt1Gc8610	Bt1G4788	304-1	g131646	97	132	1.30E-07	31	14	PHOSPHORIBOSYLFORMYLGLYCINAMIDINE SYNTHASE II (FGAM SYNTHASE II) [Bacillus subtilis]
2297	Bt1Gc8617	Bt1G4789	1-208	g584815	103	127	2.60E-08	40	41	ATP SYNTHASE B CHAIN [Bacillus subtilis]
2298	Bt1Gc8619	Bt1G4790	397-1	g585920	138	207	2.40E-15	42	11	DNA-DIRECTED RNA POLYMERASE BETA CHAIN (TRANSCRIPTASE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT) [Bacillus subtilis]
2299	Bt1Gc8627	Bt1G4791	428-1	g1666507	221	244	1.10E-20	40	41	(U61226) RfbB [Leptospira interrogans]
2300	Bt1Gc8646	Bt1G4792	406-1	g482286	163	220	1.30E-16	40	9	DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain (version 2) - Bacillus subtilis []
2301	Bt1Gc8654	Bt1G4793	1-659	g2498389	124	228	5.70E-18	28	30	PROBABLE FERRIC REDUCTASE TRANSMEMBRANE

Table 1

SEQ ID NO	Contig Id	Gene Id	Position	NCBI gi	aaf_ nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
										COMPONENT [Candida albicans]
2302	Bt1Gc8641	Bt1G4794	1-200	g2634202	186	189	7.10E-15	53	35	(Z99113) similar to alkaline phosphatase [Bacillus subtilis]
2303	Bt1Gc8660	Bt1G4795	1-409	g282384	132	125	2.10E-07	29	41	quinol oxidase aa3-600 chain II - Bacillus subtilis [Bacillus subtilis]
2304	Bt1Gc8656	Bt1G4796	324-1	g115950	204	269	2.40E-23	51	32	GLUCOSE-RESISTANCE AMYLASE REGULATOR (CATABOLITE CONTROL PROTEIN) [Bacillus subtilis]
2305	Bt1Gc8662	Bt1G4797	393-1	g729934	68	117	3.00E-07	37	57	SIGNAL PEPTIDASE I (SPASE I) (LEADER PEPTIDASE I) [Bacillus caldolyticus]
2306	Bt1Gc8664	Bt1G4798	3-431	g2072720	111	158	1.20E-10	32	29	(Z95121) hypothetical protein Rv3253c [Mycobacterium tuberculosis]
2307	Bt1Gc8670	Bt1G4799	1-206	g2226222	72	108	1.10E-05	34	26	(Y14082) hypothetical protein [Bacillus subtilis]
2308	Bt1Gc8675	Bt1G4800	349-1	g2739435	232	267	4.30E-22	49	16	(U86377) (p)ppGpp synthetase [Bacillus subtilis]
2309	Bt1Gc8679	Bt1G4801	353-32	g1168595	71	110	6.90E-06	35	38	ATP SYNTHASE GAMMA CHAIN [Bacillus stearothermophilus]
2310	Bt1Gc8692	Bt1G4802	1-374	g2633908	115	150	6.40E-10	34	29	(Z99112) similar to acetylornithine deacetylase [Bacillus subtilis]

**SEQ ID NO:**

A sequential SEQ ID NO: is assigned to each contig or singleton and the SEQ ID NO: corresponds to that set forth in the sequence listing.

**Contig ID**

Contigs or singletons are assigned an arbitrary contig ID. Contigs are assembled according to the procedure set forth in Example 2.

**Gene ID**

Refers to an arbitrarily assigned Gene ID number.

**Position**

If the first numeral under the position heading is lower than the second numeral, it designates the nucleotide position which forms part of the codon that encodes the N-most terminal amino acid of the coding sequence of the *B. thuringiensis* protein or fragment thereof. If the first numeral under the position heading is higher than that found in the corresponding second position it designates the nucleotide position which forms part of the codon that encodes the C-most terminal amino acid of the coding sequence of the *B. thuringiensis* protein or fragment thereof. In cases where the first numeral is higher than its corresponding second numeral, the *B. thuringiensis* protein or fragment thereof is encoded by the complement of the sequence set forth in the sequence listing.

Table 1

**NCBI gi**

Each sequence in the GenBank public database is arbitrarily assigned a unique NCBI gi (National Center for Biotechnology Information GenBank Identifier) number. In this table, the NCBI gi number which is associated (in the same row) with a given contig or singleton refers to the particular GenBank sequence which is the best match for that sequence.

**aat\_nap score**

The aat\_nap score is reported by the nap program in the aat package. It is an alignment score in which each match and mismatch is scored based on the BLOSUM62 scoring matrix.

**Blastp-Prob**

The entries in the "Blastp-Prob" column refer to the probability that such matches occur by chance.

**BlastP Score**

Each entry in the "BlastP Score" column of the table refers to the BLASTP score that is generated by sequence comparison of the designated clone with the designated GenBank sequence.

**% Iden**

The entries in the "%Iden" column of the table refer to the percentage of identically matched nucleotides (or residues) that exist along the length of that portion of the sequences which is aligned by the BLAST comparison to generate the statistical scores presented

**% cvrg**

The %coverage is the percent of hit sequence length that matches to the query sequence (%coverage = (match length / hit total length) x 100).

**NCBI gi description**

The "NCBIgi desc" column provides a description of the NCBIgi referenced in the "NCBIgi" column.

The following tables, Table 2 through Table 5, are offered by way of illustration and not by way of limitation. It is to be understood that the present invention is not limited to the particular proteins or polypeptides or particular coding nucleotide sequences listed in Table 2 through 5.

5

Table 2. Sigma Factor Homologs

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
184	Bt1G227	g548832	114	113	1.80E-13	29	76	RNA POLYMERASE SIGMA-H FACTOR (SIGMA-30) [ <i>Pseudomonas aeruginosa</i> ]
337	Bt1G411	g141373	173	177	1.30E-13	44	99	SIGMA-K FACTOR PROCESSING REGULATORY PROTEIN BOFA (BYPASS-OF-FORESPORE PROTEIN) [ <i>Bacillus subtilis</i> ]
499	Bt1G627	g133475	1164	1175	2.30E-119	89	100	RNA POLYMERASE SIGMA-G FACTOR (STAGE III SPORULATION PROTEIN G) [ <i>Bacillus subtilis</i> ]
499	Bt1G628	g133289	755	719	4.90E-71	96	68	RNA POLYMERASE SIGMA-35 FACTOR PRECURSOR [ <i>Bacillus thuringiensis</i> ]
1076	Bt1G1739	g133282	1169	1169	1.00E-118	100	100	RNA POLYMERASE SIGMA-28 FACTOR PRECURSOR [ <i>Bacillus thuringiensis</i> ]
1108	Bt1G1857	g282367	639	672	4.70E-66	76	77	transcription initiation factor sigma H - <i>Bacillus megaterium</i> [ <i>Bacillus megaterium</i> ]
1179	Bt1G2157	g133466	1246	1273	9.60E-130	84	79	RNA POLYMERASE SIGMA FACTOR RPOD (SIGMA-A) (SIGMA-43) [ <i>Bacillus subtilis</i> ]
1189	Bt1G2201	g134763	167	119	8.20E-07	58	17	SPORULATION SIGMA-E FACTOR PROCESSING PEPTIDASE (STAGE II SPORULATION PROTEIN GA) [ <i>Bacillus subtilis</i> ]
1205	Bt1G2294	g3287912	431	431	1.60E-40	72	99	ANTI-SIGMA F FACTOR ANTAGONIST (STAGE II SPORULATION PROTEIN AA) [ <i>Bacillus coagulans</i> ]
1205	Bt1G2295	g134757	591	591	1.80E-57	79	99	ANTI-SIGMA F FACTOR (STAGE II SPORULATION PROTEIN AB) [ <i>Bacillus licheniformis</i> ]
1205	Bt1G2296	g464690	1071	1023	3.00E-103	83	100	RNA POLYMERASE SIGMA-F FACTOR (STAGE II SPORULATION PROTEIN AC) (SPORULATION SIGMA FACTOR) [ <i>Bacillus megaterium</i> ]
1227	Bt1G2414	g2633716	565	587	4.80E-57	49	100	(Z99111) similar to RNA polymerase sigma factor [ <i>Bacillus subtilis</i> ]

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1248	Bt1G2533	g1731060	2059	2038	8.30E-211	61	97	PUTATIVE SIGMA L-DEPENDENT TRANSCRIPTIONAL REGULATOR IN MMGE-BFMBAA INTERGENIC REGIÓN [Bacillus subtilis]
1262	Bt1G2627	g133292	835	855	1.90E-85	42	100	RNA POLYMERASE SIGMA-54 FACTOR [Bacillus subtilis]
1274	Bt1G2697	g1941918	240	283	7.80E-25	35	99	(X93081) sigma F/sigma G transcribed gene [Bacillus subtilis]
1328	Bt1G3067	g3386359	711	721	3.00E-71	53	100	(AF074855) RNA polymerase sigma B [Listeria monocytogenes]
1357	Bt1G3331	g3024615	156	208	6.90E-17	25	99	RNA POLYMERASE SIGMA FACTOR SIGV [Bacillus subtilis]
1357	Bt1G3332	g548832	157	172	4.50E-13	26	99	RNA POLYMERASE SIGMA-H FACTOR (SIGMA-30) [Pseudomonas aeruginosa]
1364	Bt1G3404	g3688548	86	185	1.90E-14	28	81	(AJ010320) RNA polymerase sigma factor [Streptomyces coelicolor]
1721	Bt1G4125	g133466	99	114	4.10E-06	33	23	RNA POLYMERASE SIGMA FACTOR RPOD (SIGMA-A) (SIGMA-43) [Bacillus subtilis]
1764	Bt1G4171	g133481	172	203	2.30E-16	50	71	POSSIBLE RNA POLYMERASE SIGMA- G FACTOR (ORF3) [Bacillus thuringiensis]
2212	Bt1G4696	g133475	272	298	2.00E-26	63	36	RNA POLYMERASE SIGMA-G FACTOR (STAGE III SPORULATION PROTEIN G) [Bacillus subtilis]



**SEQID NO:** A sequential SEQ ID NO: is assigned to each contig or singleton and the SEQ ID NO: corresponds to that set forth in the sequence listing.

**Gene ID:** Refers to an arbitrarily assigned Gene ID number.

**NCBI gi:** Each sequence in the GenBank public database is arbitrarily assigned a unique NCBI gi (National Center for Biotechnology Information GenBank Identifier) number. In this table, the NCBI gi number which is associated (in the same row) with a given contig or singleton refers to the particular GenBank sequence which is the best match for that sequence.

**aat\_nap score:** The aat\_nap score is reported by the nap program in the aat package. It is an alignment score in which each match and mismatch is scored based on the BLOSUM62 scoring matrix.

**Blastp-Prob:** The entries in the "Blastp-Prob" column refer to the probability that such matches occur by chance.

**BlastP Score:** Each entry in the "BlastP Score" column of the table refers to the BLASTP score that is generated by sequence comparison of the designated clone with the designated GenBank sequence.

**% Iden:** The entries in the "%Iden" column of the table refer to the percentage of identically matched nucleotides (or residues) that exist along the length of that portion of the sequences which is aligned by the BLAST comparison to generate the statistical scores presented

**% cvrg:** The %coverage is the percent of hit sequence length that matches to the query sequence (%coverage = (match length / hit total length) x 100).

**NCBI gi description:** The "NCBIgidesc" column provides a description of the NCBIgi referenced in the "NCBIgi" column.

**Table 3. Transposases, Integrases, and Transposons**

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Iden	% Cvrg	NCBI gi description
2	Bt1G2	g2497382	150	145	2.30E-09	34	33	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
64	Bt1G88	g3005554	93	140	4.30E-09	22	36	(AF047044) putative transposase [Anabaena PCC7120]
226	Bt1G276	g549114	374	410	2.70E-38	80	19	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS231F [Bacillus thuringiensis]
379	Bt1G464	g2497400	184	229	4.10E-19	38	60	HYPOTHETICAL TRANSPOSASE-LIKE PROTEIN HI1721 [Haemophilus influenzae Rd]
383	Bt1G468	g2497400	171	211	3.30E-17	36	57	HYPOTHETICAL TRANSPOSASE-LIKE PROTEIN HI1721 [Haemophilus influenzae Rd]
387	Bt1G475	g549113	337	296	7.50E-26	72	21	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS231E [Bacillus thuringiensis]
410	Bt1G502	g2127290	57	123	1.80E-07	34	32	transposase (insertion sequence IS231) - Bacillus thuringiensis [Bacillus thuringiensis]
416	Bt1G512	g2497382	537	570	4.40E-57	85	29	TRANSPOSASE FOR INSERTION

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
								SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
546	Bt1G691	g2497382	606	641	9.00E-63	94	31	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
555	Bt1G707	g3005554	86	113	4.00E-06	29	22	(AF047044) putative transposase [Anabaena PCC7120]
603	Bt1G782	g3005554	79	142	2.60E-09	29	28	(AF047044) putative transposase [Anabaena PCC7120]
642	Bt1G840	g3426013	288	362	3.30E-33	28	85	(AB016803) transposase [Deinococcus radiodurans]
644	Bt1G843	g1789981	110	135	1.10E-08	44	20	(AE000433) IS150 putative transposase [Escherichia coli]
660	Bt1G864	g2497382	109	125	3.40E-07	96	6	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
691	Bt1G912	g1694898	597	635	1.70E-61	36	37	(Y09450) transposase [Pseudomonas putida]
691	Bt1G913	g79972	194	299	2.70E-25	26	34	transposase tnpA - Enterococcus faecalis plasmid pAD2 transposon Tn917 [Transposon Tn917]
781	Bt1G1068	g1789981	276	299	1.60E-26	54	35	(AE000433) IS150 putative transposase [Escherichia coli]
799	Bt1G1102	g1694898	358	411	2.80E-37	30	29	(Y09450) transposase [Pseudomonas putida]
980	Bt1G1481	g1749770	308	351	4.90E-32	46	52	(Y09946) transposase [Bacillus thuringiensis]
1002	Bt1G1538	g2497400	183	231	2.50E-19	40	57	HYPOTHETICAL TRANSPOSASE-LIKE PROTEIN HI1721 [Haemophilus influenzae Rd]
1045	Bt1G1654	g2497382	912	882	2.60E-88	91	45	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1072	Bt1G1724	g3005554	105	162	1.60E-11	24	41	(AF047044) putative transposase [Anabaena PCC7120]
1098	Bt1G1819	g2497382	424	470	1.20E-44	94	23	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1190	Bt1G2215	g2497382	286	292	9.10E-26	76	18	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1207	Bt1G2302	g2497382	532	548	6.50E-53	99	25	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1214	Bt1G2345	g2497382	528	544	1.70E-52	99	24	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1252	Bt1G2554	g136144	2449	2407	6.60E-250	99	100	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS231C [Bacillus thuringiensis]

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1273	Bt1G2694	g3005554	100	138	7.20E-09	24	37	(AF047044) putative transposase [Anabaena PCC7120]
1275	Bt1G2705	g135956	459	718	6.30E-71	29	100	TRANSPOSASE B (TRANSPOSON TN554) [Staphylococcus aureus]
1305	Bt1G2905	g2098612	342	433	9.90E-41	38	100	(U66614) putative transposase [Marinococcus halophilus]
1317	Bt1G2979	g1789981	387	406	7.20E-38	52	50	(AE000433) IS150 putative transposase [Escherichia coli]
1330	Bt1G3082	g3005554	93	139	5.60E-09	25	32	(AF047044) putative transposase [Anabaena PCC7120]
1340	Bt1G3161	g2497382	2231	2231	2.90E-231	100	100	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1353	Bt1G3278	g136144	316	340	7.10E-31	94	15	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS231C [Bacillus thuringiensis]
1354	Bt1G3289	g3426013	330	510	6.90E-49	30	94	(AB016803) transposase [Deinococcus radiodurans]
1362	Bt1G3378	g2497382	616	622	9.30E-61	95	29	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1378	Bt1G3590	g1749770	280	308	1.70E-27	52	41	(Y09946) transposase [Bacillus thuringiensis]
1378	Bt1G3595	g2497400	187	229	4.10E-19	38	57	HYPOTHETICAL TRANSPOSASE-LIKE PROTEIN HI1721 [Haemophilus influenzae Rd]
1380	Bt1G3625	g1789981	796	806	3.00E-80	52	100	(AE000433) IS150 putative transposase [Escherichia coli]
1383	Bt1G3662	g1789981	852	809	1.40E-80	54	100	(AE000433) IS150 putative transposase [Escherichia coli]
1386	Bt1G3684	g3005554	117	208	7.90E-17	20	72	(AF047044) putative transposase [Anabaena PCC7120]
1386	Bt1G3687	g3218350	169	240	2.80E-20	26	100	(AL023861) putative IS element transposase [Streptomyces coelicolor]
1388	Bt1G3704	g3005554	97	135	1.50E-08	24	36	(AF047044) putative transposase [Anabaena PCC7120]
1391	Bt1G3745	g1749770	735	735	9.90E-73	51	100	(Y09946) transposase [Bacillus thuringiensis]
1392	Bt1G3762	g136144	2443	2424	1.00E-251	100	100	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS231C [Bacillus thuringiensis]
1549	Bt1G3942	g2497400	92	120	2.30E-07	28	46	HYPOTHETICAL TRANSPOSASE-LIKE PROTEIN HI1721 [Haemophilus influenzae Rd]
1573	Bt1G3970	g2497382	153	179	4.60E-13	43	21	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1611	Bt1G4009	g2497382	129	139	1.00E-08	49	17	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1698	Bt1G4102	g2497382	177	220	1.50E-17	44	24	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1725	Bt1G4130	g2497382	115	159	6.90E-11	52	14	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1739	Bt1G4145	g2497382	270	276	7.20E-24	49	29	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1804	Bt1G4211	g2497382	213	248	1.10E-20	56	22	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
1869	Bt1G4277	g549114	93	161	5.10E-11	28	26	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS231F [Bacillus thuringiensis]
1902	Bt1G4312	g1789981	258	286	3.80E-25	35	58	(AE000433) IS150 putative transposase [Escherichia coli]
1965	Bt1G4386	g2497382	218	203	1.10E-15	36	40	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
2041	Bt1G4473	g136144	266	307	4.10E-27	66	20	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS231C [Bacillus thuringiensis]
2049	Bt1G4484	g2497382	119	154	2.40E-10	35	22	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
2130	Bt1G4607	g2497382	130	155	1.90E-10	56	13	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
2135	Bt1G4612	g2497382	174	229	1.50E-18	49	27	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
2153	Bt1G4631	g2497382	130	187	6.20E-14	61	13	TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS232 [Insertion sequence IS232]
102	Bt1G134	g586206	178	178	1.90E-13	67	16	D-ALANYL-D-ALANINE CARBOXYPEPTIDASE (DD- PEPTIDASE) (DD- CARBOXYPEPTIDASE) [Transposon Tn1546]
1340	Bt1G3160	g141450	74	136	1.40E-08	32	29	HYPOTHETICAL 37.1 KD PROTEIN IN TRANSPOSON TN4556 [Transposon Tn4556]
1795	Bt1G4202	g1196998	220	208	2.50E-16	55	24	(J01829) unknown protein [Transposon Tn10]
1797	Bt1G4204	g1196998	160	180	3.00E-13	51	18	(J01829) unknown protein [Transposon Tn10]
1989	Bt1G4413	g4309763	475	494	6.20E-46	75	10	(AC006217) putative retrotransposon polyprotein [Arabidopsis thaliana]
2055	Bt1G4494	g479357	235	251	1.60E-20	65	14	hypothetical protein 612 - maize transposon MuA2 [Zea mays]

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvr	NCBI gi description
2057	Bt1G4496	g2130141	254	184	4.20E-13	51	14	mudrA protein - maize transposon MuDR [Zea mays]
2248	Bt1G4735	g1196998	108	118	1.70E-06	61	11	(J01829) unknown protein [Transposon Tn10]
14	Bt1G21	g1710383	584	606	4.60E-59	72	53	PROBABLE INTEGRASE/RECOMBINASE RIPX [Bacillus subtilis]
296	Bt1G357	g4098413	95	173	1.20E-12	32	32	(U77495) putative integrase [Leuconostoc oenos bacteriophage 10MC]
722	Bt1G961	g4490997	364	442	1.10E-41	29	100	(AL035707) putative integrase [Streptomyces coelicolor]
834	Bt1G1164	g1881291	294	381	3.20E-35	29	95	(AB001488) PROBABLE INTEGRASE. [Bacillus subtilis]
834	Bt1G1165	g166159	321	335	2.40E-30	30	95	(M34832) integrase (int) [Bacteriophage phi-11]
999	Bt1G1529	g1926326	693	724	1.40E-71	39	100	(X98106) integrase [Bacteriophage phigle]

**SEQID NO:** A sequential SEQ ID NO: is assigned to each contig or singleton and the SEQ ID NO: corresponds to that set forth in the sequence listing.

**Gene ID:** Refers to an arbitrarily assigned Gene ID number.

**NCBI gi:** Each sequence in the GenBank public database is arbitrarily assigned a unique NCBI gi (National Center for Biotechnology Information GenBank Identifier) number. In this table, the NCBI gi number which is associated (in the same row) with a given contig or singleton refers to the particular GenBank sequence which is the best match for that sequence.

**aat nap score:** The aat nap score is reported by the nap program in the aat package. It is an alignment score in which each match and mismatch is scored based on the BLOSUM62 scoring matrix.

**Blastp-Prob:** The entries in the "Blastp-Prob" column refer to the probability that such matches occur by chance.

**BlastP Score:** Each entry in the "BlastP Score" column of the table refers to the BLASTP score that is generated by sequence comparison of the designated clone with the designated GenBank sequence.

**% Ident:** The entries in the "%Ident" column of the table refer to the percentage of identically matched nucleotides (or residues) that exist along the length of that portion of the sequences which is aligned by the BLAST comparison to generate the statistical scores presented

**% cvrg:** The %coverage is the percent of hit sequence length that matches to the query sequence (%coverage = (match length / hit total length) x 100).

**NCBI gi description:** The "NCBIGidesc" column provides a description of the NCBIgi referenced in the "NCBIgi" column.

**Table 4. Antibiotic, Chemical, and Heavy Metal Resistance**

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvr	NCBI gi description
-----------	---------	---------	---------------	--------------	--------------	---------	-------	---------------------

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
33	Bt1G51	g266515	250	273	2.10E-22	29	19	MULTIDRUG RESISTANCE-LIKE ATP-BINDING PROTEIN MDL [Escherichia coli]
98	Bt1G130	g3861147	104	114	4.80E-06	33	17	(AJ235272) BICYCLOMYCIN RESISTANCE PROTEIN (bcr1) [Rickettsia prowazekii]
145	Bt1G181	g1174516	499	526	1.30E-49	55	17	ISOLEUCYL-TRNA SYNTHETASE, MUPIROCIN RESISTANT (ISOLEUCINE--TRNA LIGASE) (ILERS) (MUPIROCIN RESISTANCE PROTEIN) [Staphylococcus aureus]
162	Bt1G201	g1684651	432	314	4.00E-28	37	56	(Z82987) unknown similar to quinolon resistance protein NorA [Bacillus subtilis]
180	Bt1G222	g1945096	183	192	1.70E-14	54	19	(D88802) S. lividans chloramphenicol resistance protein; P31141 (492) transmembrane [Bacillus subtilis]
204	Bt1G249	g136472	237	284	6.10E-25	33	99	ACETYLTRANSFERASE (TABTOXIN RESISTANCE PROTEIN) [Pseudomonas syringae]
275	Bt1G332	g2108269	116	133	6.10E-09	41	41	(X92868) mercuric resistance operon regulatory protein [Bacillus subtilis]
298	Bt1G359	g1684651	383	298	2.00E-26	38	48	(Z82987) unknown similar to quinolon resistance protein NorA [Bacillus subtilis]
361	Bt1G441	g3861147	199	215	4.30E-17	27	60	(AJ235272) BICYCLOMYCIN RESISTANCE PROTEIN (bcr1) [Rickettsia prowazekii]
397	Bt1G486	g1705428	604	605	5.90E-59	59	84	BACITRACIN RESISTANCE PROTEIN (PUTATIVE UNDECAPRENOL KINASE) [Escherichia coli]
421	Bt1G519	g1174634	93	168	2.10E-12	26	60	TELLURITE RESISTANCE PROTEIN TEHB HOMOLOG [Haemophilus influenzae Rd]
423	Bt1G522	g2633162	329	245	7.60E-26	25	83	(Z99108) similar to multidrug resistance protein [Bacillus subtilis]
579	Bt1G744	g1684651	638	349	9.00E-55	33	100	(Z82987) unknown similar to quinolon resistance protein NorA [Bacillus subtilis]
613	Bt1G795	g2632985	1127	958	2.30E-96	37	66	(Z99107) similar to acriflavin resistance protein [Bacillus subtilis]
624	Bt1G809	g2688027	183	189	1.70E-13	22	27	(AE001125) acriflavine resistance protein (acrB) [Borrelia burgdorferi]
692	Bt1G915	g2145816	117	144	4.20E-10	28	99	bacitracin resistance protein homolog bacA - Mycobacterium leprae [Mycobacterium leprae]
726	Bt1G970	g2500765	396	404	1.20E-37	28	100	SENSOR PROTEIN VANSB (VANCOMYCIN B-TYPE RESISTANCE PROTEIN VANSB) (VANCOMYCIN HISTIDINE PROTEIN KINASE) [Enterococcus faecalis]
862	Bt1G1220	g2827439	1774	1774	7.80E-183	78	100	(AF043609) aluminum resistance protein [Arthrobacter viscosus]

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
930	Bt1G1354	g2499116	335	371	3.70E-34	29	100	VANCOMYCIN B-TYPE RESISTANCE PROTEIN VANW [Enterococcus faecalis]
950	Bt1G1408	g1277135	374	374	1.80E-34	41	42	(U50978) kanamycin/gentamycin- resistance protein [Cloning vector pFW13]
986	Bt1G1495	g399406	246	212	2.60E-17	25	85	DAUNORUBICIN RESISTANCE TRANSMEMBRANE PROTEIN [Streptomyces peucetius]
995	Bt1G1520	g1174516	1843	1846	1.80E-190	58	53	ISOLEUCYL-TRNA SYNTHETASE, MUPIROCIN RESISTANT (ISOLEUCINE--TRNA LIGASE) (ILERS) (MUPIROCIN RESISTANCE PROTEIN) [Staphylococcus aureus]
1005	Bt1G1550	g2879772	110	145	3.30E-10	37	45	(Y07640) putative mercury resistance operon regulatory protein (MerR) [Listeria monocytogenes]
1023	Bt1G1597	g1945096	952	778	2.70E-77	46	100	(D88802) S. lividans chloramphenicol resistance protein; P31141 (492) transmembrane [Bacillus subtilis]
1130	Bt1G1938	g1652918	412	300	1.20E-26	28	100	(D90909) quinolone resistance protein NorA [Synechocystis sp.]
1188	Bt1G2199	g585375	1044	1054	1.60E-106	70	100	DIMETHYLADENOSINE TRANSFERASE (S- ADENOSYLMETHIONINE-6-N', N'- ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RRNA DIMETHYLASE) (HIGH LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN DIMETHYLTRANSFERASE) [Bacillus subtilis]
1190	Bt1G2211	g2633434	931	787	3.10E-78	42	100	(Z99109) similar to multidrug resistance protein [Bacillus subtilis]
1208	Bt1G2308	g2634168	444	462	8.40E-44	59	99	(Z99113) similar to fosfomycin resistance protein [Bacillus subtilis]
1226	Bt1G2406	g1705428	713	561	2.70E-54	55	100	BACITRACIN RESISTANCE PROTEIN (PUTATIVE UNDECAPRENOL KINASE) [Escherichia coli]
1227	Bt1G2407	g1881228	159	173	2.40E-12	36	20	(AB001488) SIMILAR TO ENZYMES WHICH ACT VIA AN ATP- DEPENDENT COVALENT BINDING OF AMP TO THEIR SUBSTRATE. [Bacillus subtilis]
1240	Bt1G2497	g1881342	392	441	1.40E-41	34	100	(AB001488) SIMILAR TO BACILLUS CEREUS ZWITTERMICIN A- RESISTANCE GENE. [Bacillus subtilis]
1246	Bt1G2519	g461637	168	276	4.30E-24	23	100	MULTIDRUG RESISTANCE PROTEIN 1 (MULTIDRUG-EFFLUX TRANSPORTER 1) [Bacillus subtilis]
1246	Bt1G2520	g728970	180	293	6.80E-26	21	100	MULTIDRUG RESISTANCE PROTEIN 2 (MULTIDRUG-EFFLUX TRANSPORTER 2) [Bacillus subtilis]

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
1257	Bt1G2596	g728970	1488	1194	2.30E-121	72	100	MULTIDRUG RESISTANCE PROTEIN 2 (MULTIDRUG-EFFLUX TRANSPORTER 2) [Bacillus subtilis]
1272	Bt1G2692	g1881374	892	804	4.80E-80	46	100	(AB001488) SIMILAR TO BICYCLOMYCIN RESISTANCE PROTEIN. [Bacillus subtilis]
1302	Bt1G2884	g115950	1180	1191	4.70E-121	75	92	GLUCOSE-RESISTANCE AMYLASE REGULATOR (CATABOLITE CONTROL PROTEIN) [Bacillus subtilis]
1339	Bt1G3158	g2632985	3348	2412	3.20E-256	63	100	(Z99107) similar to acriflavin resistance protein [Bacillus subtilis]
1355	Bt1G3297	g2500765	301	395	1.10E-36	29	67	SENSOR PROTEIN VANSB (VANCOMYCIN B-TYPE RESISTANCE PROTEIN VANSB) (VANCOMYCIN HISTIDINE PROTEIN KINASE) [Enterococcus faecalis]
1374	Bt1G3544	g4914624	1071	920	2.50E-92	57	92	(AJ009627) multidrug resistance transporter [Listeria monocytogenes]
1393	Bt1G3769	g2769708	955	852	3.90E-85	37	100	(U82085) pristinamycin resistance protein VgaB [Staphylococcus aureus]
1426	Bt1G3813	g2827439	122	143	3.70E-09	49	17	(AF043609) aluminum resistance protein [Arthrobacter viscosus]
1460	Bt1G3849	g994737	104	125	1.30E-07	52	20	(M18327) kanamycin resistance protein [cloning vectors]
1471	Bt1G3861	g2633162	161	156	1.40E-10	25	44	(Z99108) similar to multidrug resistance protein [Bacillus subtilis]
1526	Bt1G3918	g585375	126	176	2.70E-13	35	45	DIMETHYLADENOSINE TRANSFERASE (S-ADENOSYLMETHIONINE-6-N', N'-ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RRNA DIMETHYLASE) (HIGH LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN DIMETHYLTRANSFERASE) [Bacillus subtilis]
1854	Bt1G4261	g1945096	446	406	7.20E-38	59	35	(D88802) S. lividans chloramphenicol resistance protein; P31141 (492) transmembrane [Bacillus subtilis]
1914	Bt1G4327	g2668553	130	68	0.85	37	6	(U62929) multidrug resistance protein 1 [Filobasidiella neoformans]
1923	Bt1G4339	g115950	133	153	1.80E-10	55	18	GLUCOSE-RESISTANCE AMYLASE REGULATOR (CATABOLITE CONTROL PROTEIN) [Bacillus subtilis]
2151	Bt1G4629	g3925779	238	181	5.00E-13	29	31	(AL034353) putative major facilitator family multi-drug resistance protein [Schizosaccharomyces pombe]
2179	Bt1G4659	g2632985	98	107	9.60E-05	30	11	(Z99107) similar to acriflavin resistance protein [Bacillus subtilis]
2211	Bt1G4695	g4416482	101	115	2.40E-06	38	24	(AF125999) daunorubicin resistance protein A [Mycobacterium avium]



SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
2304	Bt1G4796	g115950	204	269	2.40E-23	51	32	GLUCOSE-RESISTANCE AMYLASE REGULATOR (CATABOLITE CONTROL PROTEIN) [Bacillus subtilis]

**SEQID NO:** A sequential SEQ ID NO: is assigned to each contig or singleton and the SEQ ID NO: corresponds to that set forth in the sequence listing.

**Gene ID:** Refers to an arbitrarily assigned Gene ID number.

**NCBI gi:** Each sequence in the GenBank public database is arbitrarily assigned a unique NCBI gi (National Center for Biotechnology Information GenBank Identifier) number. In this table, the NCBI gi number which is associated (in the same row) with a given contig or singleton refers to the particular GenBank sequence which is the best match for that sequence.

**aat\_nap score:** The aat\_nap score is reported by the nap program in the aat package. It is an alignment score in which each match and mismatch is scored based on the BLOSUM62 scoring matrix.

**Blastp-Prob:** The entries in the "Blastp-Prob" column refer to the probability that such matches occur by chance.

**BlastP Score:** Each entry in the "BlastP Score" column of the table refers to the BLASTP score that is generated by sequence comparison of the designated clone with the designated GenBank sequence.

**% Ident:** The entries in the "%Ident" column of the table refer to the percentage of identically matched nucleotides (or residues) that exist along the length of that portion of the sequences which is aligned by the BLAST comparison to generate the statistical scores presented

**% cvrg:** The %coverage is the percent of hit sequence length that matches to the query sequence (%coverage = (match length / hit total length) x 100).

**NCBI gi description:** The "NCBIgi desc" column provides a description of the NCBIgi referenced in the "NCBIgi" column.

**Table 5. Toxins and Toxin Homologs**

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	BlastP Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
73	Bt1G100	g419952	125	135	1.30E-09	29	13	alpha-latroinsectotoxin precursor - black widow spider (fragment) [Latrodectus tredecimguttatus]
212	Bt1G260	g2507017	330	263	1.00E-22	68	29	HEMOLYSIN BL BINDING COMPONENT PRECURSOR (ENTEROTOXIN 40 KD SUBUNIT) [Bacillus cereus]
316	Bt1G386	g2507017	990	1027	1.10E-103	68	78	HEMOLYSIN BL BINDING COMPONENT PRECURSOR (ENTEROTOXIN 40 KD SUBUNIT) [Bacillus cereus]
410	Bt1G503	g2507017	126	156	1.10E-10	60	13	HEMOLYSIN BL BINDING COMPONENT PRECURSOR (ENTEROTOXIN 40 KD SUBUNIT)

SEQ ID NO	Gene Id	NCBI gi	aat nap Score	<u>BlastP</u> Score	BlastP- Prob	% Ident	% Cvrg	NCBI gi description
479	Bt1G597	g2507017	1829	1732	2.20E-178	98	98	[Bacillus cereus] HEMOLYSIN BL BINDING COMPONENT PRECURSOR (ENTEROTOXIN 40 KD SUBUNIT)
694	Bt1G917	g1665720	1623	1644	4.70E-169	95	93	[Bacillus cereus] (D17312) diarrheal toxin [Bacillus cereus]
1552	Bt1G3948	g2507017	210	200	1.50E-15	51	25	HEMOLYSIN BL BINDING COMPONENT PRECURSOR (ENTEROTOXIN 40 KD SUBUNIT)
2056	Bt1G4495	g97193	96	105	9.60E-05	51	6	[Bacillus cereus] leukotoxin B - Pasteurella haemolytica []

**SEQID NO:** A sequential SEQ ID NO: is assigned to each contig or singleton and the SEQ ID NO: corresponds to that set forth in the sequence listing.

**Gene ID:** Refers to an arbitrarily assigned Gene ID number.

5 **NCBI gi:** Each sequence in the GenBank public database is arbitrarily assigned a unique NCBI gi (National Center for Biotechnology Information GenBank Identifier) number. In this table, the NCBI gi number which is associated (in the same row) with a given contig or singleton refers to the particular GenBank sequence which is the best match for that sequence.

10 **aat\_nap score:** The aat\_nap score is reported by the nap program in the aat package. It is an alignment score in which each match and mismatch is scored based on the BLOSUM62 scoring matrix.

**Blastp-Prob:** The entries in the "Blastp-Prob" column refer to the probability that such matches occur by chance.

**BlastP Score:** Each entry in the "BlastP Score" column of the table refers to the BLASTP score that is generated by sequence comparison of the designated clone with the designated GenBank sequence.

15 **% Iden:** The entries in the "%Iden" column of the table refer to the percentage of identically matched nucleotides (or residues) that exist along the length of that portion of the sequences which is aligned by the BLAST comparison to generate the statistical scores presented

**% cvrg:** The %coverage is the percent of hit sequence length that matches to the query sequence (%coverage = (match length / hit total length) x 100).

20 **NCBI gi description:** The "NCBIgidesc" column provides a description of the NCBIgi referenced in the "NCBIgi" column.

**What is claimed is:**

1. An isolated nucleic acid molecule having a first nucleotide sequence, wherein: (1) the first nucleotide sequence hybridizes under stringent conditions to a second nucleotide sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283, wherein the hybridizing portion of the second nucleotide sequence is at least 50 nucleotides in length; (2) the first nucleotide sequence is a portion of the third nucleotide sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283 and is at least 50 nucleotides in length; (3) the first nucleotide sequence encodes a *B. thuringiensis* polypeptide or protein, wherein the *B. thuringiensis* polypeptide or protein is any polypeptide or protein set forth in Table 1; or (4) the first nucleotide sequence is the complement of (1), (2) or (3).
2. The isolated nucleic acid molecule of claim 1, wherein: (1) the first nucleotide sequence hybridized under stringent conditions to the second nucleotide sequence, wherein the hybridizing portion of the second nucleotide sequence is at least 100 nucleotides in length; or (2) the first nucleotide sequence is the complement of (1).
3. The isolated nucleic acid molecule of claim 2, wherein the hybridizing portion of the second nucleotide sequence is at least 200 nucleotides in length.
4. The isolated nucleic acid molecule of claim 1, wherein: (1) the first nucleotide sequence is the portion of the third nucleotide sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283 and is at least 50 nucleotides in length; or (2) the first nucleotide sequence is the complement of (1).
5. The isolated nucleic acid molecule of claim 4, wherein the portion of the third nucleotide sequence is a regulatory sequence.
6. The isolated nucleic acid molecule of claim 4, wherein the portion of the third nucleotide sequence is a promoter or partial promoter sequence.
7. The isolated nucleic acid molecule of claim 1, wherein: (1) the first nucleotide sequence encodes a *B. thuringiensis* polypeptide or protein, wherein the *B. thuringiensis* polypeptide or protein is

any polypeptide or protein set forth in Table 1; or (2) the first nucleotide sequence is the complement of (1).

8. The isolated nucleic acid molecule of claim 7, wherein the first nucleotide sequence encodes a *B. thuringiensis* polypeptide or protein which is any one set forth in Table 1.
9. The isolated nucleic acid molecule of claim 8, wherein the *B. thuringiensis* polypeptide or protein is an insect inhibitory protein or polypeptide homologue.
10. The isolated nucleic acid molecule of claim 8, wherein the *B. thuringiensis* protein or polypeptide is a sigma factor homologue.
11. The isolated nucleic acid molecule of claim 8, wherein the amino acid sequence of the *B. thuringiensis* polypeptide or protein is selected from the group consisting of SEQ ID Nos: 33, 98, 145, 162, 180, 204, 275, 298, 361, 397, 421, 423, 579, 613, 624, 692, 726, 862, 930, 950, 986, 995, 1005, 1023, 1130, 1188, 1190, 1208, 1226, 1227, 1240, 1246, 1246, 1257, 1272, 1302, 1339, 1355, 1374, 1393, 1426, 1460, 1471, 1526, 1854, 1914, 1923, 2151, 2179, 2211, and 2304.
12. The isolated nucleic acid molecule of claim 8, wherein the *B. thuringiensis* polypeptide or protein is a transposase homologue.
13. The isolated nucleic acid molecule of claim 8, wherein the *B. thuringiensis* polypeptide or protein is an integrase homologue.
14. The isolated nucleic acid molecule of claim 8, wherein the *B. thuringiensis* polypeptide or protein is selected from the group consisting of 2, 64, 226, 379, 383, 387, 410, 416, 546, 555, 603, 642, 644, 660, 691, 691, 781, 799, 980, 1002, 1045, 1072, 1098, 1190, 1207, 1214, 1252, 1273, 1275, 1305, 1317, 1330, 1340, 1353, 1354, 1362, 1378, 1378, 1380, 1383, 1386, 1386, 1388, 1391, 1392, 1549, 1573, 1611, 1698, 1725, 1739, 1804, 1869, 1902, 1965, 2041, 2049, 2130, 2135, 2153, 102, 1340, 1795, 1797, 1989, 2055, 2057, 2248, 14, 296, 722, 834, 834, and 999.
15. The isolated nucleic acid molecule of claim 8, wherein the *B. thuringiensis* protein or polypeptide has antibiotic, heavy metal, or other chemical resistance properties.

16. The isolated nucleic acid molecule of claim 8, wherein said nucleic acid molecule further comprises a promoter or partial promoter region.
17. A substantially purified *B. thuringiensis* polypeptide or protein comprising an amino acid sequence, wherein the amino acid sequence is defined as follows: (1) the amino acid sequence is encoded by a first nucleotide sequence which specifically hybridizes to the complement of a second nucleotide sequence selected from the group consisting of SEQ ID No: 1 to SEQ ID No: 8283; or (2) the amino acid sequence is encoded by a third nucleotide sequence that is at least 50% identical to an open reading frame set forth in Table 1.
18. The substantially purified *B. thuringiensis* polypeptide or protein of claim 17, wherein the *B. thuringiensis* polypeptide or protein is any protein or polypeptide set forth in Table 1.
19. The substantially purified *B. thuringiensis* protein or polypeptide of claim 18, wherein the *B. thuringiensis* protein or polypeptide is an insect inhibitory protein.
20. The substantially purified *B. thuringiensis* protein or polypeptide of claim 18, wherein the *B. thuringiensis* protein or polypeptide is a sigma factor homologue.
21. The substantially purified *B. thuringiensis* protein or polypeptide of claim 18, wherein the *B. thuringiensis* protein or polypeptide is a transposase homologue.
22. The substantially purified *B. thuringiensis* protein or polypeptide of claim 18 wherein the *B. thuringiensis* protein or polypeptide is an integrase homologue.
23. The substantially purified *B. thuringiensis* protein or polypeptide of claim 18, wherein the *B. thuringiensis* protein or polypeptide is a toxin or toxin homologue.
24. The substantially purified *B. thuringiensis* protein or polypeptide of claim 18, wherein the *B. thuringiensis* protein or polypeptide has antibiotic, heavy metal, or other chemical resistance properties.

25. A transformed cell comprising an exogenous nucleic acid molecule which comprises:
  - a) an exogenous promoter region which functions in said cell to cause the production of an mRNA molecule; which is operably linked to
  - b) a structural nucleotide sequence, wherein said structural nucleotide sequence encodes a *B. thuringiensis* protein or polypeptide which is any protein or polypeptide set forth in Table 1; which is operably linked to
  - c) a 3' non-translated sequence that functions in said cell to cause termination of transcription.
26. The transformed cell according to claim 25, wherein the *B. thuringiensis* protein or polypeptide is an insect inhibitory protein or polypeptide.
27. The transformed cell according to claim 25, wherein said *B. thuringiensis* protein or polypeptide is a sigma factor homologue.
28. The transformed cell according to claim 25, wherein said *B. thuringiensis* protein or polypeptide is a transposase homologue.
29. The transformed cell according to claim 25, wherein said *B. thuringiensis* protein or polypeptide is an integrase.
30. The transformed cell according to claim 25, wherein said *B. thuringiensis* protein or polypeptide is a protein homologue having antibiotic resistance properties.
31. The transformed cell according to claim 25, wherein said *B. thuringiensis* protein or polypeptide is a toxin or toxin homologue.
32. The transformed cell according to claim 26, wherein said cell is selected from the group consisting of a bacterial cell, a plant cell, an algal cell, a mammalian cell, an insect cell and a fungal cell.
33. A transformed plant comprising an exogenous nucleic acid which comprises:

- a) an exogenous promoter region which functions in a plant cell to cause the production of an mRNA molecule; which is operably linked to
  - b) a structural nucleotide sequence encoding a polypeptide or protein set forth in Table 1; which is operably linked to
  - c) a 3' non-translated sequence that functions in said plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to the 3' end of said mRNA molecule.
34. The transformed plant of claim 33, wherein the polypeptide or protein is an insect inhibitory polypeptide or protein.
35. The transformed plant according to claim 33, wherein said plant is a monocot or a dicot plant.
36. A computer readable medium having recorded thereon one or more nucleotide sequences, wherein each of the nucleotide sequences is selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 8283 or complements thereof.
37. The computer readable medium according to claim 36, wherein each of the nucleotide sequences or complements thereof encodes a *B. thuringiensis* protein or polypeptide.
38. A method for generating a transgenic plant comprising the steps of: a) introducing into the genome of the plant an exogenous nucleic acid, wherein the exogenous nucleic acid comprises in the 5' to 3' direction i) a promoter that functions in the cells of said plant, said promoter operably linked to; ii) a structural nucleotide sequence encoding a polypeptide or protein set forth in Table 1, said structural nucleic acid sequence operably linked to; iii) a 3' non-translated nucleic acid sequence that functions in said cells of said plant to cause transcriptional termination; b) obtaining transformed plant cells containing the nucleic acid sequence of step (a); and c) regenerating from said transformed plant cells a transformed plant in which said polypeptide or protein is overexpressed.
39. The method of claim 38, wherein the protein or polypeptide is an insect inhibitory polypeptide or protein.

40. A method for identifying one or more genes encoding insect inhibitory proteins in the sequences of one or more plasmids of a *Bacillus thuringiensis*, the method comprising the steps of:
- a) isolating and purifying plasmid DNA;
  - b) constructing a DNA library from the isolated and purified plasmid DNA;
  - c) sequencing the DNA library to obtain a set of plasmid DNA sequences;
  - d) comparing the set of DNA sequences with a set of chromosomal DNA sequences, wherein the set of chromosomal DNA sequences comprises the group consisting of SEQ ID No: 1 through SEQ ID No: 8283;
  - e) identify common sequences, which are identified both in the set of plasmid DNA sequences and in the set of chromosomal DNA sequences;
  - f) subtracting the common sequences from the set of plasmid DNA sequences to obtain a subtracted set of plasmid DNA sequences;
  - g) assembling the subtracted set of DNA sequences to contigs and sequences;
  - h) determining open reading frames in the contigs and sequences; and
  - i) identifying one or more genes encoding insect inhibitory proteins in the sequences of one or more plasmids of said *Bacillus thuringiensis*.
41. A method for identifying plasmid DNA sequences of a *Bacillus* species, the method comprising the steps of:
- a) identifying a *Bacillus* species strain which does not contain plasmid DNA;
  - b) generating a library of chromosomal genomic DNA from said *Bacillus* species strain which does not contain plasmid DNA;
  - c) obtaining the nucleotide sequence of said chromosomal genomic DNA;
  - d) identifying a *Bacillus* species strain which contains plasmid DNA;
  - e) generating a library of said *Bacillus* species plasmid DNA;
  - f) obtaining the nucleotide sequence of said plasmid DNA;
  - g) subtracting any common sequences identified in the plasmid DNA which are also identified in the chromosomal genomic DNA; and
  - h) constructing contigs and sequences of said plasmid DNA;
- wherein said contigs and sequences comprise the plasmid DNA sequence of said *Bacillus* species.



42. The method according to claim 41 wherein said *Bacillus* species is selected from the group consisting of *Bacillus thuringiensis*, *Bacillus subtilis*, *Bacillus cereus*, and *Bacillus anthracis*.
43. The method according to claim 42 wherein said *Bacillus* species is *Bacillus thuringiensis*.
44. The method of claim 41 wherein the nucleotide sequence of said chromosomal genomic DNA comprises the sequences selected from the group consisting of SEQ ID NO:1 through SEQ ID NO:8283.
45. The method of claim 41 wherein said *Bacillus* species strain contains less than two naturally occurring plasmids.
46. The method according to claim 45 wherein said *Bacillus* species strain is EG 10650.
47. The method of claim 45 wherein said less than two naturally occurring plasmids is selectively tagged with an identifiable marker gene.
48. The method according to claim 46 wherein said marker gene is selected from the group consisting of an antibiotic resistance gene, a gene encoding an essential metabolic or catabolic protein or functional homologue thereof, a gene conferring bioluminescence properties, and a gene encoding an enzyme which catalyzes the metabolism of a substrate which imparts a colored product deposited on or within the *Bacillus* species strain.
49. An isolated and purified nucleic acid sequence selected from the group consisting of SEQ ID NO:1 through SEQ ID NO:8283
50. An isolated and purified *Bacillus thuringiensis* protein or polypeptide which is selected from the group consisting of proteins or polypeptides identified in Table 1.

### Abstract

The present invention relates to nucleic acid sequences from *Bacillus thuringiensis* and, in particular, to genomic DNA sequences. The invention encompasses nucleic acid molecules present in non-coding regions as well as nucleic acid molecules that encode proteins, fragments of proteins, tRNA's, fragments of tRNA's, rRNA's and fragments of rRNA's. In addition, proteins and fragments of proteins so encoded and antibodies capable of binding the proteins are encompassed by the present invention. The invention also relates to methods of using the disclosed nucleic acid molecules, proteins, fragments of proteins, RNA's, and antibodies, for example, for gene identification and analysis, and preparation of constructs.